Attention-Induced Trading and Returns: Evidence from Robinhood Users

Brad M. Barber
Graduate School of Management
UC Davis

Xing Huang
Olin Business School
Washington University in St. Louis

Terrance Odean
Haas School of Business
University of California, Berkeley

Chris Schwarz
Merage School of Business
UC Irvine

First Draft: October 18, 2020

July 2021

We appreciate the comments of Azi Ben-Rephael (discussant), Charles Jones (discussant), Michaela Pagel (discussant), Ivo Welch, and seminar and conference participants at the University of Central Florida, the University of Missouri, Erasmus University Rotterdam, Maastricht University, Ohio State University, Q Group, 3rd Virtual QES NLP and Machine Learning in Investment Management Conference, FSU SunTrust Beach Conference, SFS Cavalcade, China Meeting of the Econometric Society, and the Western Finance Association meetings.

Attention-Induced Trading and Returns: Evidence from Robinhood Users

Abstract

We study the influence of financial innovation by fintech brokerages on individual investors' trading and stock prices. Using data from Robinhood, we find that Robinhood investors engage in more attention-induced trading than other retail investors. For example, Robinhood outages disproportionately reduce trading in high-attention stocks. While this evidence is consistent with Robinhood attracting relatively inexperienced investors, we show that it can also be partially driven by the app's unique features. Consistent with models of attention-induced trading, intense buying by Robinhood users forecast negative returns. Average 20-day abnormal returns are -4.7% for the top stocks purchased each day.

Over the past half century, investor trading has changed significantly. Decades ago, retail investors traded via phone only during market hours, paying heavy commissions to do so. The 1990s brought about online trading and significantly lower commissions. More recently, the fintech brokerage Robinhood brought about even more changes. Robinhood was the first brokerage to offer commission-free trading on a convenient, simple, and engaging mobile app. In contrast to the dramatic changes in the investment landscape, the changes in investment psychology are likely less dramatic. Do these changes in the investment landscape alter individual investors' trading behavior?

On one hand, the lack of commissions and simplicity may reduce the costs and barriers to investing in the stock market. Even small costs can reduce stock market participation for less wealthy households (Vissing-Jorgensen, 2002). Thus, the simplicity of the Robinhood app and similar fintech applications may increase stock market participation.

On the other hand, simplicity is not problem free. To its app, Robinhood "added features to make investing more like a game. New members were given a free share of stock, but only after they scratched off images that looked like a lottery ticket."[1] New and inexperienced investors may find these features appealing. However, some believe that Robinhood over-emphasizes the fun of trading at the cost of sound investment practices. In December 2020, Massachusetts state regulators filed a complaint against Robinhood citing its "aggressive tactics to attract inexperienced investors" and "use of strategies such as gamification to encourage and entice continuous and repetitive use of its trading application."[2] Indeed, Robinhood users are unusually active. In the first quarter of 2020, Robinhood users "traded nine times as many shares as E-Trade customers, and 40 times as many shares as Charles Schwab customers, per dollar in the average customer account in the most recent quarter."[3] Thus while Robinhood's innovations may have had a positive influence on market participation (and Robinhood's customer acquisition), their influence on trading behavior is an open question. With these issues in mind, we study the behavior of Robinhood users using data on aggregate Robinhood user changes at the stock-day level from May 2018 to August 2020.

We first conjecture Robinhood users are more likely to be influenced by attention than other investors. Half of Robinhood users are first-time investors,[4] who are unlikely to have developed their own clear criteria for buying a stock. Inexperienced stock investors are more heavily influenced by attention (Seasholes and Wu, 2007) and by biases that lead to return chasing (Greenwood and Nagel, 2009). With turnover rates many times higher than those of other brokerage firms, Robinhood users are more likely to trade speculatively. As a result, a smaller proportion of their trading is motivated by non-speculative reasons

[1] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html
[2] https://www.sec.state.ma.us/sct/current/sctrobinhood/MSD-Robinhood-Financial-LLC-Complaint-E-2020-0047.pdf
[3] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html
[4] https://blog.robinhood.com/news/2020/5/4/robinhood-raises-280-million-in-series-f-funding-led-by-sequoia

such as saving for retirement, meeting liquidity needs, harvesting tax losses, or rebalancing their portfolio. The higher rate of speculative trading by Robinhood users increases the potential for attention-driven trading.

If Robinhood users are more likely than other investors to be influenced by attention, their purchase behavior is more likely to be correlated; that is, they are more likely to herd than other investors. This is exactly what we find. We document that 35% of net buying by Robinhood users is concentrated in 10 stocks compared to 24% of net buying by the general population of retail investors. We then analyze herding episodes by Robinhood users. We define a herding episode as a day when the number of Robinhood users owning a particular stock increases dramatically. In our primary analysis, we focus on the top 0.5% of positive user changes as a percent of prior day user count each day. This represents about 10 herding episodes per day or almost 5,000 episodes over the 26-month sample period. We show that these herding episodes are predicted by attention measures (e.g., recent investor interest, extreme returns, or unusual volume). Finally, we show that during Robinhood outages retail trading drops more in high-attention stocks than in other stocks relative to periods with no outage. The evidence from Robinhood outages provides strong evidence that Robinhood users are more likely to engage in attention-induced trading than other retail investors.

The simplicity of Robinhood's app is likely to guide investor attention for three reasons. First, the app prominently displays lists of stocks in an environment relatively free of complex information. For example, besides basic market information, Robinhood only provides five charting indicators, while TD Ameritrade provides 489.[5] This streamlined and simplified interface likely guides the choices of Robinhood users. Second, the Robinhood app makes it very easy to place trades and the reduction of frictions increases trading (Barber and Odean, 2002). Third, the simplification of information on the Robinhood app is likely to provide cognitive ease to investors, leading them to rely more on their intuition and less on critical thinking, or more on System 1 thinking and less on System 2 (Kahneman 2011). Of course, the Robinhood app is not the only channel through which the attention of Robinhood users becomes focused on the same subset of stocks. For example, many Robinhood users share information and opinions on online forums such as Reddit's WallStreetBets.[6]

To identify the effect of the app on Robinhood users, we focus on the "Top Mover" list, which lists only 20 stocks and changes every day (and throughout each day). Crucially, this list displays stocks with the largest absolute percentage price changes from the previous day close. In contrast, many websites provide separate lists of stocks with the largest daily gains and losses (e.g., Yahoo! Finance Gainers and Yahoo! Finance Losers), and on these sites top gainers tend to be more prominently displayed. Moreover,

[5] https://www.stockbrokers.com/compare/robinhood-vs-tdameritrade
[6] https://www.wsj.com/articles/the-real-force-driving-the-gamestop-amc-blackberry-revolution-11611965586?page=1

Google search volume suggests that investors are about twice as likely to look for stocks with same day gains than those with same day losses.[7] Thus, if the app itself is driving Robinhood users' trading, we would expect Robinhood traders to buy both gainers and losers heavily, while other retail investors will tend to buy gainers. This is precisely what we find: Robinhood users are drawn to trading both extreme gainers and losers, whereas other retail investors prefer to buy extreme gainers rather than losers. While prior work documents that investors buy extreme winners and losers (Barber and Odean, 2008), our evidence indicates the Robinhood app affects the intensity of this behavior because of the unique way Robinhood displays the "Top Mover" list.

We provide additional evidence that the "Top Mover" list influences Robinhood user buying behavior by exploiting another unique feature of the list. Robinhood requires stocks to be above $300 million in market capitalization to be displayed in the top movers list. We use a sharp regression discontinuity design to show that Robinhood users are more likely to buy stocks with market capitalizations between $300 and $350 million that were in the top twenty stocks when sorting on absolute return than stocks with similar absolute returns but market capitalizations between $250 and $300 million. Thus stocks that just miss making the list due to market capitalization below the $300 million cutoff do not get the increase in users associated with being on the "Top Movers" list.

Models of attention-induced trading and returns predict that periods of intense buying will be followed by negative abnormal returns (e.g., Barber and Odean, 2008; Pedersen 2021). We conjecture that the concentrated buying of Robinhood users, who are susceptible to attention-induced trading, provides an unusually strong setting to identify the return effects of attention-induced trading. In our final set of analyses, we focus on this return prediction and document large negative abnormal returns following Robinhood herding episodes. Specifically, the top 0.5% of stocks bought every day lose about 4.7% over the subsequent month.

The magnitude of the negative abnormal returns increases dramatically as we identify fewer, but more intense herding episodes. To systematically analyze the relation between the herding intensity and price reversal, we analyze stocks with a minimum of 100 Robinhood users and identify different sets of herding episodes by varying the daily percentage increase in users holding the stock from 10% to 750%. At a 10% increase in users, we observe over 20,000 herding episodes; at a 750% increase, we observe 45 episodes. The large negative abnormal returns in the month following these herding events grow from a statistically significant -1.8% when we require a 10% increase in users (> 20,000 events) to an extremely large and statistically significant -19.6% when we require a 750% increase in users (45 events).

[7] Google trends indicates the phrase "top gainers today" ("top stock gainers today") is searched more than twice as much as "top losers today" ("top stock losers today") for the five years beginning January 24, 2016.

The negative returns that follow purchase herding by Robinhood users are not simply inventory-based reversals as modeled in Jagadeesh and Titman (1995) and documented around earnings announcements in So and Wang (2014). Attention-induced trading can also cause return reversals when investors intensely buy stocks with strong recent returns. However, the negative return following herding episodes is not completely limited to return reversals. First, in multivariate analyses, in which we control for return reversals, the negative abnormal returns following herding episodes remain large and statistically significant. Second and more importantly, we observe negative returns following a day when we observe a surge in Robinhood users that is preceded by an overnight drop in the stock's price, perhaps because aggressive Robinhood user buying slows the stock's response to negative news (Barber and Odean, 2008). The negative returns we document following purchase herding by Robinhood users are also not driven by the bid-ask spread since they persist when we use quote midpoints to calculate returns.

Given the relatively small size of Robinhood, one might question whether Robinhood users have the potential to influence market prices. Robinhood has $81 billion in assets under custody,[8] far less than E*TRADE, $600 billion, TD Ameritrade, $1.3 trillion, and Charles Schwab, $3.8 trillion.[9] However, trades, not passive positions, move prices. There are a lot of Robinhood users: 13 million as of May of 2020 compared to 12.7 million at Schwab and 5.5 million at E*TRADE at the end of 2019.[10] And, as noted above, Robinhood users are extremely active traders. In June of 2020, Robinhood users averaged 4.3 million revenue trades per day (Daily Average Revenue Trades or DARTs), more than E*TRADE, 1.1 million, TD Ameritrade, 3.8 million, Charles Schwab, 1.8 million, Interactive Brokers, 1.9 million, or Fidelity, 1.4 million.[11] Thus Robinhood users accounted for roughly 30% of the daily trades from the largest brokerage firms catering to retail investors and have the potential to move prices. Furthermore, as noted above, Robinhood trades may be a good proxy for the actions of other attention-motivated traders who herd in the same stocks.

Two additional findings indicate the negative returns we observe are caused, at least in part, by the trading of Robinhood users and other attention-motivated retail investors. First, we expect the influence of investors to be most pronounced in small cap stocks. In the cross-section, we find negative returns following Robinhood herding events for stocks with market caps under $1 billion, but not for stocks with market caps over $1 billion. Second, we expect bigger effects during periods with heightened retail trading. In time series, retail trading has increased significantly at Robinhood and elsewhere in the COVID-19 period (i.e.,

---

[8] See page 26 of Robinhood's SEC Form S1 IPO filing (https://www.sec.gov/Archives/edgar/data/1783879/000162828021013318/robinhoods-1.htm)
[9] https://www.businessofapps.com/data/robinhood-statistics/
[10] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html
[11] https://www.bloomberg.com/news/articles/2020-08-10/robinhood-blows-past-rivals-in-record-year-for-retail-investing. Fidelity's daily trades are for all of 2020, not just June (https://www.barrons.com/articles/fidelitys-trading-volume-surged-in-the-pandemic-but-its-struggling-to-boost-revenue-51614702735). Note that brokerages exercise some discretion in measuring DARTs.

after March 13, 2020) and the negative return effect following Robinhood herding events is more pronounced in the COVID-19 period. Irrespective of whether the negative returns we document result from the trading of Robinhood users and other attention-motivated retail investors, these negative returns lead to trading losses for many investors.

Savvy investors might exploit the trading patterns and predictably negative returns that we document. To profit in response to Robinhood herding events, an investor would sell the stock short (or, equivalently, purchase put options that the option seller would hedge by shorting). Thus, if investors are exploiting Robinhood user herding, we would expect to see increased short interest around Robinhood user herding events. Indeed, we do find a marked increase in short selling for stocks involved in Robinhood herding events even after controlling for returns and news.

While the "Top Mover" list and, possibly, other features of the Robinhood app influence Robinhood users, many drivers of attention will affect Robinhood users and other retail investors similarly. We show that this is the case using a measure of retail investor herding developed by Barber, Lin, and Odean (2021). BLO classify herding events as TAQ retail trades in the top quintile of retail standardized order imbalance and top decile of abnormal retail volume. For each day in our Robinhood sample, we modify the BLO methodology to identify the same number of TAQ herding events as we identify using our Robinhood herding measure. 73% (3,567 of 4,884) of the Robinhood and the BLO herding events are identical. Like Robinhood herding events, these events are, on average, followed by negative returns. However, Robinhood herding events that are accompanied by retail selling on TAQ also earn negative returns and the subsequent negative returns for non-Robinhood herding events are significantly smaller in magnitude than those that are related to Robinhood events. These results show that the actions of Robinhood users provide a good proxy for attention-induced trading.

Our study is of particular interest given the unique dataset of the retail investors (Robinhood users) that we analyze. To our knowledge, four papers use the same dataset. Welch (2020) analyzes the holdings and performance of Robinhood users. He concludes Robinhood users principally held stocks with large persistent past volume and do not underperform with respect to standard academic benchmark models.[12] In a JP Morgan report, Cheng, Murphy, & Kolanovic (2020) show that users are drawn to stocks that attract investor attention and that changes in stock popularity predict returns. Unlike these studies, we document poor returns following extreme attention-driven herding events by Robinhood users. Ozik, Sadka, and Shen (2020) use the Robintrack data to analyze the sharp increase in retail trading and the effect on bid-ask spreads during the pandemic period. They find an increase in trading of stocks with COVID-19 related media coverage, which they attribute to an attention-grabbing effect. They also document the increase in

[12] We too find that the aggregate performance of Robinhood users is not reliably different from zero using standard asset pricing technology. See Internet Appendix Table A1.

retail trading generally lowered stock bid-ask spreads and price impact of trades. Along the same line of inquiry, Eaton, Green, Roseman, & Wu (2021) use Robinhood outages to study the effect of retail trading on market quality and find that these negative shocks to Robinhood participation reduce market volatility and improve liquidity. In contrast to these studies, we study the attention-induced trading of Robinhood investors and show herding episodes by Robinhood investors reliably predict negative returns.

In summary, we provide two main contributions to the academic literature. First, we present evidence that the Robinhood app influences investors behavior. Specifically, we show that the prominently featured "Top Mover" list (which displays only 20 stocks, sorts stocks on absolute (rather than signed) percentage returns, and changes regularly) affects Robinhood users. This finding fits into the emerging literature that emphasizes the display of information affects investor behavior. Changes in the display of price information affects investors willingness to sell winners versus losers in individual stocks and mutual funds (Frydman & Wang, 2020; Loos, Meyer, & Pagel, 2020). News that investors consume about stocks often confirms our prior beliefs (Cookson, Engelberg, and Mullins, 2021), and its prominence affects the incorporation of information (Fedyk 2019). Displaying return performance for index funds can lead investors to prefer high-fee funds (Choi, Laibson & Madrian, 2010), prominently featuring expense information can lead investors to prefer low-fee funds (Kronlund, Pool, Sialm & Stefanescu, 2020), and the prominence of mutual fund lists affects investors fund choices (Kaniel and Parham, 2017; Hong, Lu, and Pan 2019).[13] Our results, and this emerging literature, indicate that disclosure alone is not sufficient to assure good investor outcomes; how information is displayed can both help and hurt investors. And, while the recent literature on complexity in finance emphasizes its dark side (Carlin 2009), our results suggest simple user interfaces are not necessarily the solution to problems that arise from complexity; both complexity and simplicity can lead investors astray.

Second, we contribute to the literature that documents the effects of attention-induced trading on returns. Using Robinhood trading as a proxy for attention-induced herding episodes, we find strong support for the return predictions of attention-based models of trading. Specifically, we link episodes of intense buying by Robinhood users to negative returns following the herding episodes. Our focus on trading, rather than events, is distinct from the extant literature that focuses on events like Jim Kramer's stock recommendations (Engelberg, Sasseville, & Williams, 2012; Keasler & McNeil, 2010; Bolster, Trahan & Venkateswaran, 2012), the WSJ Dartboard Column (Liang, 1999; Barber & Loeffler, 1993), Google stock searches (Da, Engelberg, & Gao, 2011; Da, Hua, Hung, & Peng, 2020), and repeat news stories (Tetlock, 2011). Perhaps as a result of focusing on trading behavior rather than events, the magnitude of the return

[13] Da, Larrain, Sialm, and Tessada (2018) find that recommendations of an advisory firm followed by many Chilean pension investors generate correlated fund flows and influence market returns. These attention-induced effects on the active choice of mutual funds are different from the stickiness of default options (e.g., Cronqvist and Thaler, 2004), which might result from inertia or a view that defaults are an implicit recommendation.

reversals we document are much larger and more widespread than those documented in prior studies.[14] Moreover, the extant literature documents negative returns subsequent to positive return events, we find negative returns following intense buying that coincides with or follows negative returns.

1 Data and Methods

In this section, we describe the main Robintrack dataset that keeps track of how many Robinhood users hold a particular stock over time and our methods for identifying extreme herding events by Robinhood users.

1.1 Robintrack Data

The primary dataset for our analysis comes from the Robintrack website (https://robintrack.net/), which scrapes stock popularity data from Robinhood between May 2, 2018, and August 13, 2020.[15] Robinhood discontinued the reporting of stock popularity data on August 13, 2020. The Robintrack dataset contains repeated cross-sectional snapshots of user counts for individual securities (e.g., 645,535 Robinhood users held Apple stock at 3:46 pm ET on August 3, 2020).[16] Our main results include all Robintrack securities since we do not have strong priors about what types of securities will experience herding events.[17]

We merge the Robintrack data to CRSP and TAQ data by using the ticker on Robintrack. The CRSP database provides daily returns, closing and opening prices, closing bid-ask spreads, and market capitalization. We use the TAQ database to identify retail buys and sells using the Boehmer, Jones, Zhang, & Zhang (BJZZ, 2020) algorithm. The BJZZ algorithm relies on the observation that retail trades often receive price improvement in fractions of a penny and are routed to a FINRA trade reporting facility (TRF). Thus, the BJZZ algorithm identifies retail buys as trades reported to a FINRA trade reporting facility (exchange code "D" in TAQ) with fractional penny prices between \$0.006 and \$0.01; retail sells are trades reported to a FINRA TRF with fraction penny prices between \$0.00 and \$0.004.[18]

In Figure 1, Panel A, we see the total number of Robinhood user-stock positions grew from about 5 million at the beginning of our sample period to more than 42 million at the end. In May 2020, Robinhood

---

[14] See Internet Appendix Table A2 for a summary of these studies. The biggest magnitude of return reversal is -4.6% for 39 events over three years (Barber & Loeffler, 1993). In a widely cited study, Da, Engelberg, & Gao (2011) fail to find robust evidence of price reversals following spikes in Google search volume.

[15] About 11 dates during the sample period are missing user data, four in January 2019 and seven in January 2020. For 16 dates on which we observe users, no observations were recorded between 2 and 4 pm ET.

[16] The Robintrack data are generally reported every hour at approximately 45 minutes after the hour. The data from Robinhood has some lag. Thus, the user count at 3:46 on Robintrack for Apple is from sometime before 3:46. Based on some analysis of open data, the likely lag is between 30 and 45 minutes. The Robinhood App appears to update data every 15 minutes.

[17] The results are similar for common stocks and other securities, though US common stocks represent 70% of all herding episodes, stocks with non-US headquarters 13%, and ADRs 10%.

[18] We also use TAQ to calculate returns in July and August 2020 since CRSP data were not available through August 2020 at the writing of this draft.

reported having 13 million users, which translates into about 3 stock positions per user.[19] The red line in the figure denotes the date on which the COVID-19 national emergency was declared in the US (March 13, 2020); there is a clear increase in Robinhood users after this date. In Figure 1, Panel B, we plot the total number of TAQ retail trades for comparison. Retail trading also increased during the pandemic period. Of course, some Robinhood trades are part of these retail trades.

The Robintrack data does not allow us to identify individual trades, but it does allow us to analyze changes in user positions in a particular stock. The analogue to this Robinhood user change variable in TAQ is net retail buying in a stock. In Figure 1, Panel C, we plot the five-day moving average of the daily sum across stocks of the absolute value of Robinhood user changes (green line) and five-day moving average of the daily sum across stocks of the absolute value of TAQ net buying, i.e., number of retail buys minus the number of retail sells (blue line). Both measures of trading activity follow similar trajectories with a marked increase in the pandemic period.

We present additional summary statistics in Table 1, Panel A, across stock-day observations. The main variable, $users\_close$, measures the total number of users in a stock prior to the close of trading (4 pm ET) but after 2 pm on the same day. The key variables in our later analysis of herding events are based on the daily changes in $users\_close$ ($userchg$) or the ratio of $users\_close$ on consecutive days ($userratio$). For descriptive purposes, we also report $users\_last$, which is the last reported user count for a stock on each day (regardless of the time of reporting).

The mean stock has a bit more than 2,000 users, though the median user count is 160. User changes are generally small; the interquartile range of *userchg* is 2 and of *userratio* is 0.01. Similarly, TAQ net buying is generally quite small, with an interquartile range of -7 to +10. In Table 1, Panel B, we present descriptive statistics across days. The average day has 7,211 stock holdings and just under 15 million user-positions.

1.2 Herding Events

While we find that user changes are generally small, there are a number of extreme user change events. These extreme events likely occur because Robinhood users are new to markets and more willing to speculate. They are also likely a good proxy for the behavior of investors who are unduly influenced by attention-grabbing events. To catalog these herding events, we identify stocks with an increase in users (i.e., $userratio(t) > 1$) and at least 100 users entering the day (i.e., $users\_close(t-1) \geq 100$). Among these stocks, we sort stocks based on the day t $userratio$ and identify the top 0.5% of stocks as Robinhood herding stocks, which we denote with the indicator variable rh_herd. This results in the identification of 4,884 herding events (about 9 per day on average).

[19] See "Robinhood Has Lured Young Traders, Sometimes with Devastating Results" in NY Times, July 8, 2020. (https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html)

In Table 2, we present descriptive statistics on the herding events across stock-day observations. The average stock in these episodes has about 2,500 users and experiences an increase in users of 1,100. Of particular interest is the return on the stock on the day of these episodes, which is on average 14%, with most of the return occurring at the open of trading; the mean opening return ($openret$) is 11%. Despite these large positive mean daily returns, we also observe a number of stocks (about 1/3) with large negative returns on the day of these herding events. As we point out later, the appeal of large negative stocks may be partially a function of the Robinhood app, which highlights "Top Movers" for the day based on absolute rather than signed returns. Thus, unlike many stock lists that focus on the most positive movers for the day, the Robinhood app focuses its users' attention on stocks with extreme returns. We tend to observe a large retail order imbalance in TAQ on these days as well, which is not surprising since Robinhood trades are a subset of TAQ trades.

2 Attention and Stock Selection

In the first part of our analysis, we document that Robinhood users show excessively concentrated trading activities, compared with the general population of retail investors measured by TAQ dataset. We then show attention measures strongly predict Robinhood herding episodes and use the model to forecast the probability of herding episodes for individual stocks. To convincingly show Robinhood users are particularly prone to trading in attention-grabbing stocks, we exploit Robinhood trading outages and document sharper drops in retail volume for stocks with a high probability of a herding episode during these outages. In a final analysis, we show unique features of the Robinhood platform predict the trading of Robinhood users more strongly than other retail investors which suggests the app itself guides user decisions.

2.1 The Concentration of Buying versus Selling

In theory, attention-induced trading should predominantly affect purchase rather than sale decisions. Retail investors can choose to buy any stock that captures their attention but can only sell stocks that they own (unless they sell short, which is relatively uncommon among retail investors and not possible on the Robinhood platform). We expect attention motivated trading to be common among Robinhood users. To test this conjecture, we compare the concentration of buying activity to the concentration of selling activity for Robinhood users. We also anticipate the concentration of buying will be greater for Robinhood users than the general retail investor population. While attention certainly affects the general population of retail investors, we expect other motives to play a greater role in their trading decisions (e.g., trade to rebalance, harvest tax losses, diversify, or save/consume rather than attention-motivated trading). This is especially true since half of Robinhood users are new investors, who are more subject to attention biases (Seasholes and Wu, 2007).

To empirically investigate these issues, we first identify the 10 stocks with the most Robinhood-buying activity on each day (i.e., largest Robinhood user changes). We then calculate the concentration of buying among these 10 stocks as the total number of new users for these stocks divided by sum of user increases for all stocks with an increase in users. Similarly, we identify the 10 stocks with the most TAQ retail buying (i.e., largest net buying based on number of retail buys minus number of retail sells). We then calculate the concentration of buying among these 10 stocks as the total number of net buys for these stocks divided by sum of net buying for all stocks with net buying (i.e., a positive order imbalance). This calculation is repeated on each day, yielding a time-series of daily measures of buying concentration for Robinhood and TAQ retail investors. There is an analogous calculation for negative user changes on Robinhood and TAQ net selling.

In Figure 2, we present the mean concentration of buying (Panel A) and selling (Panel B) for Robinhood users (green bars) and TAQ retail trades (white bars). Consistent with the idea that attention has a bigger effect on buying than selling, for both Robinhood and TAQ retail traders, the concentration of buying (Panel A) is higher than the concentration of selling (Panel B). However, concentrations of both buying and selling are stronger for Robinhood investors than the general population of retail traders.[20]

In Table 3, we summarize the mean percentage of trades observed in the top 10 stocks and also calculate mean daily Herfindahl-Hirschman (HH) indexes for buying (Panel A) and selling (Panel B). For Robinhood users, about 35% of all net buying is in the top 10 stocks; for TAQ retail trades, 24% of net buying is observed in these stocks. In contrast, for Robinhood users, 25% of selling is concentrated in the top 10 stocks; for TAQ retail trades, 14% is concentrated in these stocks. The HH indexes for buying are larger than those observed for selling for both Robinhood and TAQ, which indicates that, in general, the concentration of buying activity is higher than the concentration of selling activity for retail traders. Moreover, the HH indexes for Robinhood are greater than those observed for TAQ for both buying and selling, which indicates a higher degree of buying and selling concentration for Robinhood users.

One concern with the above analysis is that Robinhood user changes (new owners of the stock) do not map perfectly to TAQ net buying. Essentially, we assume that measures of new owners and net buying generate similar concentration statistics. To test this conjecture, we use the discount broker trade and position data of Barber and Odean (2000) to calculate two variables: daily new users for each stock (from daily positions) and daily order imbalance for each stock (from trades). The correlation between these two series is 87%; they generate concentration statistics in the broker data that differ by less than 0.5% at the

[20] This conclusion assumes that there is no bias in the BJZZ methodology that would affect concentration measures.

individual stock level. Given the differences in concentration that we observe in Table 3 and Figure 2, we conclude that the differences are not due to measurement differences and therefore Robinhood users purchase activity is indeed more concentrated than the general population of retail traders.

2.2 Attention Proxies and Robinhood Herding Events

If attention is guiding the trading decision of Robinhood investors, we expect proxies for investor attention to be strong predictors of Robinhood herding episodes. Here, we examine the relation between a collective set of attention measures and the herding episodes we study. While this analysis is interesting on its own, the model also identifies stocks that are at high risk of a herding episode, which will be useful for our analysis of retail trading during Robinhood outages. To begin, we estimate a linear probability model by regressing the extreme herding episode indicator on a set of attention measures, including extreme absolute lagged returns, lagged abnormal volume, lagged user change, lagged level of users, lagged abnormal Google search volume, lagged abnormal news coverage, and lagged earnings announcement. We process Google search volume index (SVI) following Da, Engelberg, and Gao (2011) and Niessner (2015). We construct the news coverage variable by counting the daily number of news articles written on the ticker based on data obtained from Thomson Reuters MarketPsych Indices (TRMI). All abnormal measures on day $t-1$ are computed as the logarithm of the ratio of the value on day $t-1$ to the average from day $t-21$ to $t-2$. The lagged herding indicator is also included in the regression to capture the persistence in the herding episodes. Robust standard errors are clustered by day and stock level.

Table 4 presents the results. We find persistence in the herding episodes. The coefficient on the lagged herding indicator is positive and statistically significant. A stock which is heavily bought by Robinhood investors is 10% more likely to experience another episode the next day. This is not surprising: the herding episode itself may generate discussion and attract attention through media or social media platforms and lead to additional herding the next day. Moreover, consistent with our results that Robinhood investors respond to the "Top Mover" list, we find that if a stock's absolute return is ranked in the top 20, the probability of this stock being heavily bought the next day increases by about 5%; this is statistically significant at 1% level. In addition, the other attention measures all lead to higher probability of heavy buying activities on Robinhood.

Taken together, we find that the extreme herding episodes of Robinhood investors are persistent and can be predicted by a set of attention measures. We will use the predicted value from these specifications to capture the attention-driven component of the extreme herding episodes in our analyses of Robinhood outages.

2.3 The Effect of Robinhood Outages on Retail Trading

In this section, we build on the linear probability model of the prior section to establish the importance of Robinhood user trading in attention-grabbing stocks. To do so, we exploit three unexpected trading outages on the Robinhood user platform. These outages allow us to estimate the impact of Robinhood trading on retail trading in general. More importantly, the outages allow us to demonstrate a sharper drop in retail trading for stocks that are good candidates for the herding events we study or popular among Robinhood users. This evidence supports our claim that Robinhood users are more likely to engage in attention-induced trading than other retail investors.

To identify Robinhood outages, we review the incident history on Robinhood websites. There are three outages that affected equity trading on March 2, March 3, and June 18. The most prolonged outage occurred on March 2 and lasted virtually the entire trading day (listed as under investigation at 9:38 am ET and was posted as resolved at 2:13 am ET on March 3). The next day, March 3, there was an intraday outage between 10:04 am ET (posted as under investigation) with service partially restored at 11:35 am ET and fully restored at 11:55 am ET. The third outage occurred on June 18 and began at 11:39 am ET (posted as under investigation) with improvement at 12:43 pm ET (post indicating "starting to see improvement") and resolution at 1:08 pm ET.

To estimate the economic impact of Robinhood trading, we use these outages, which are arguably exogenous events that prevent Robinhood users from trading but have no effect on retail investors who trade using other brokers.[21] To do so, we measure the proportion of retail trading relative to all trading with hourly intervals during the trading hours (i.e., 9:30 am to 4:00 pm ET, with the first interval spanning 9 am to 10 am ET). Retail trades are identified in TAQ as in Boehmer et al. (BJZZ, 2020).

In Figure 3, we show the mean proportion of retail trading for the 50 most popular stocks on Robinhood during these key outage events. Outages are depicted with red bars. In Panel A, the full day March 2 outage has the lowest percent of retail trade. In Panel B, we see that mean retail trading during the 10 am hour was low on both March 2 (full day outage) and March 3 (intraday outage). In Panel C, we see that mean retail trading during the noon hour was low on March 2 (full day outage) but high when trading resumed on Robinhood following an early outage on March 3. In Panel D, we see that mean retail trading between 11:35 am and 12:40 pm ET is low on June 18 relative to other days.

To more formally test for differences, we estimate the following regression for the day-long March 2 outage:

$$RetailPerc_{it} = a + bOutage_t + \mu_{tod} + \mu_{stock} + e_{it} , \tag{1}$$

[21] One might be concerned that unusually heavy volume caused the Robinhood outage. We do not think this is a major concern since the March 2 full-day outage has volume that ranks 12th out of the 21 days centered on March 2. March 3rd ranks 9th during the same period. June 18 is the lowest volume day during the 21 days centered on June 18.

where $RetailPerc_{it}$ is the percent of trades that are retail trades on TAQ during hour *t* for stock *i*. $Outage_t$ is an indicator variable that takes a value of one on March 2. As controls, we include time of day fixed effects (μ_{tod}) and stock fixed effects (μ_{stock}). The key coefficient estimate, *b*, measures the percentage point decline in the percentage of total trading during the Robinhood outage period. Standard errors are double clustered by day and stock.

For the intraday outage on March 3, we estimate the following regression:

$$RetailPerc_{it} = a + bOutage_t + cRepair_t + \mu_t + \mu_{tod} + \mu_{stock} + e_{it} . \quad (2)$$

For this episode, *Outage* is an indicator variable that takes a value of 1 between 10 and 11 am on March 3, 2020, and *Repair* is an indicator variable that takes a value of 1 between noon and 1 pm (the hour after systems are fully restored).

For the intraday outage on June 18, $RetailPerc_{it}$ is measured at 5-minute intervals to estimate the following regression:

$$RetailPerc_{it} = a + bOutage_t + cPartial + dRepair_t + \mu_t + \mu_{tod} + \mu_{stock} + e_{it} . \quad (3)$$

For the June episode, *Outage* takes a value of one for intervals beginning at 11:35 am to 12:35 pm, *Partial* takes a value of one for the intervals beginning 12:40 to 1:00 pm, and *Repair* takes a value of one for the intervals beginning between 1:05 and 2:00 pm.

Table 5 summarizes the results. We estimate models for all stocks, the 50 most popular Robinhood stocks, and the 50 highest attention stocks in Columns (1) to (3), respectively. To identify the 50 highest attention stocks, we use the fitted values from the linear probability model that predicts the herding Robinhood herding events (Column (3), Table 4). In Panel A, we see the full day outage on March 2 reduces trading for all stocks by 0.723 percentage points (ppt) ($p < .001$), which represents 6% of the average fraction of retail trading (12.10%) during this period. For the 50 most popular Robinhood stocks, retail trading declines by 5.227 ppt ($p < .001$), which represents 36% of the average fraction of retail trading (14.60%) for these stocks. For the 50 high attention stocks, we observe a 4.948 ppt decline in trades, which represents a 28% decline in the typical level of retail trades for these stocks (17.68%).

For the intraday outage of March 3, we observe similar patterns and similar magnitudes during the outage period. However, we also observe detectable rebounds in trading in the first hour after the outage, suggesting that the outage generated some pent-up demand to trade among retail investors. For the intraday June 18 outage, we observe similar patterns but somewhat smaller magnitudes.[22]

In summary, the analysis of outages shows that Robinhood users account for as much as 6.6% of total trades in stocks (and 1/3 of retail trades) in the 50 most popular Robinhood stocks. Perhaps more

[22] In Internet Appendix Figure A1, we show the five-minute mean trading during the 11:35 am to 12:40 pm time interval. Trading volume increases noticeably at the end of this interval. We do not know if this is random or perhaps a result of Robinhood systems being partially operational before the posted time on their website.

importantly for our purposes, the analysis also reveals Robinhood users are particularly active in high attention stocks, accounting for as much as 6.5% of total trading in these high attention stocks (and more than 1/4th of total retail trading). The latter result lends credibility to our assumption that the analysis of Robinhood trading provides a good proxy for attention-motivated trading. These magnitudes are also consistent with the June 2020 DARTs data that suggests Robinhood represents approximately 30% of retail trades.

2.4 The Robinhood User Interface and Stock Selection

One potential driver for the excessively concentrated trading on Robinhood could be the coordination of common signals. Given individuals' aversion to complexity (Umar, 2020; Oprea, 2020), Robinhood adopts a simple and sleek platform design to make the financial decision-making more cognitively accessible to investors. Its simplified interface is in striking contrast with traditional brokerage firms, which provide investors a rich set of indicators and research tools. For example, besides basic market information, Robinhood only provides five charting indicators, while TD Ameritrade provides 489.[23] Presented with a large variety of stimuli, investors using traditional investing products are likely to have heterogenous responses given the limited capacity and highly flexible allocation of human attention (Kahneman, 1973). In contrast, the reduced number of stimuli on Robinhood makes it easy for investors to focus their attention and likely generate coordinated attention-induced responses.

In this section, we analyze the influence of Robinhood's lists on investor's trading. These lists are displayed prominently on the platform and are easily accessible through the tabs under "News/Popular Lists." The two most prominently displayed lists are the "Most Popular" list and the "Top Mover" list. We focus on the "Top Mover" list because it is short and constantly changing, while the "Most Popular" list is long and largely static. "Top Mover" lists stocks with the day's largest percent gains and losses since the market close the previous day.

2.4.1 "Top Mover" Absolute Return Feature

The default sorting of the "Top Mover" list is based on the absolute returns and thus mixes top gainers and top losers.[24] This feature differs from almost all other media accounts (e.g., Wall Street Journal, Yahoo! Finance, CNBC, etc.) that also report the top movers, but separate the top gainers and top losers rather than mixing them together.

[23] https://www.stockbrokers.com/compare/robinhood-vs-tdameritrade
[24] Internet Appendix Figure A2 provides an example of the "top mover" list on October 8, 2020, as shown on the website. The initial screen shows four stocks ranked at the top by absolute returns, which includes three top gainers and one top loser.

We exploit this unusual feature of Robinhood's Top Mover list and compare the buying activity between Robinhood investors and general retail investors measured by TAQ data. Given the top gainers and top losers are displayed in the same list on Robinhood, we would expect the degree of availability for the two groups of stocks to be similar for Robinhood investors. As a result, the buying activity of Robinhood investors would not differ much between top gainers and top losers. In contrast, the top losers are less prominently displayed on other media accounts, where they are reported separately from the top gainers. Accordingly, we would expect general retail investors to respond less strongly to top losers than to top gainers.

Figure 4 presents the graphical evidence for the comparison. We measure the buying activity of Robinhood investors by intraday user change[25] and the buying behavior of retail investors by TAQ intraday retail net purchases (i.e., number of buyer-initiated retail trades minus seller-initiated retail trades).[26] In Panel A, we rank stocks based on absolute overnight returns from the market close of day $t-1$ to the market open of day t, and buying activity is measured on day t.[27] Thus buying activity is measured subsequent to ranking. The graph on the left presents how the buying activity of Robinhood users on day t varies with the rank of the 20 stocks with highest absolute value of overnight return (i.e., the 20 highest "Top Movers"); the graph on the right reports buying activity for retail investors. We plot the mean Robinhood intraday user change and TAQ net retail buying for the top 20 movers separately for stocks with positive returns (top gainers) and stocks with negative returns (top losers).

Stocks with bigger absolute price changes are bought more by both Robinhood users and retail investors. This is consistent with evidence of attention-based buying documented in Barber and Odean (2008). Robinhood investors respond similarly to top gainers and losers, while other retail investors buy top gainers much more aggressively than top losers. This is consistent with our hypothesis that the attention of Robinhood users is directed to both top winners and top losers because both appear on the Top Movers list, while the attention of other investors is more likely directed to stocks that appear on Top Gainers lists. Panel B sorts top movers on the daily close-to-close return as a robustness test and shows similar patterns.

To more formally test whether the difference is statistically significant between Robinhood intraday user change and TAQ intraday net retail buying, we estimate the following specification:

$$NetBuy_{it} = \beta_0 + \beta_1 Score_{it} + \beta_2 I_{R_{it}<0} + \beta_3 Score_{it} \times I_{R_{it}<0} + \alpha_t + \varepsilon_{it}, \quad (4)$$

[25] Since the algorithm by Boehmer et al. (2020) cannot identify retail trades at the open auction, for this analysis we exclude the user change at the open on Robinhood to make the Robinhood user change more comparable with TAQ net retail buying.

[26] Since Robinhood does not allow short selling, to make TAQ net retail buying more comparable with the Robinhood user change, we remove short trades from TAQ following Boehmer and Song (2020). The results are similar with or without short trades removed from TAQ.

[27] We do not have the actual Top Mover list as it appears on Robinhood but use the return ranks as a proxy for whether a stock is likely to appear on the list. Accordingly, we only rank stocks with market caps greater than $300 million since Robinhood only ranks stocks above this size threshold to create the Top Mover list.

where the dependent variable is the (Robinhood or TAQ) buying activity for stock i on day t, $Score_{it}$ assigns a score to each rank of top movers. For expositional ease, we assign 20 to the stock with highest absolute return, and 1 to the stock with the 20th highest absolute returns; thus scores increase with the absolute returns. $I_{R_{it}<0}$ is an indicator variable that equals one if the stock return is negative. For each day, we only include the top 20 movers into the regression. A day-fixed effect is included, and the robust standard errors are clustered on the daily level. We hypothesize that while other retail investors are less likely to buy top losers than to buy top gainers, Robinhood investors are not.

In Table 6, Columns (1) and (2) sort top mover scores on absolute overnight returns. The buying activity increases with top mover scores in both columns. This indicates that, for both Robinhood investors (Column (1)) and general retail investors (Column (2)), attention is affected by the ranks within top movers, and higher ranks make the stocks more salient.

The key difference in the buying activity of Robinhood investors and general retail investors is reflected by the coefficient on the indicator for negative returns (or, top losers). For general retail investors, the top losers garner much less buying activity than the top gainers. Within the same rank, the TAQ net retail buying decreases by 133.2 trades for a top loser versus a top winner, which is similar to the decrease associated with the rank of a top gainer dropping by ten ($\approx 133.2/13.41$). In addition, the interaction of the top rank and negative return is negative indicating the magnitude of negative return effect is larger for the more extreme returns. This pattern differs from that of Robinhood investors. If anything, Robinhood buying activity is slightly stronger for the top losers than for the top gainers and the interaction effect is positive. Columns (3) and (4) sort top mover scores on absolute daily returns and find similar results for the general population of retail traders.

An alternative explanation for the differences between Robinhood and TAQ shown in Figure 4 and Table 6 could be different combination of positive feedback traders and contrarians between Robinhood and general retail investors. In this case, investors simply respond to extreme price movements rather than the information display. There can be two distinct groups - positive feedback traders who respond positively to extreme past returns and contrarians who respond negatively to extreme past returns. If the two groups of investors are evenly distributed in Robinhood while positive feedback traders dominate contrarians among general retail investors, we would observe the patterns shown in Figure 4 and Table 6. To address this concern and provide additional evidence of the app's influence on Robinhood users, in the next section we explore another unique feature of the Top Mover list.

2.4.2 "Top Mover" Market Cap Requirement

In this section, we exploit a regression discontinuity design to further establish the causal impact of the app interface on investor's trading behavior. Specifically, Robinhood requires stocks to be above $300 million in market cap to be displayed in the top movers list.[28] This feature enables us to employ a sharp regression discontinuity design to study Robinhood investors' responses to extreme price movements of stocks with market cap around the \$300 million cutoff.[29] If the interface does not affect investors' trading behavior, there should not be any discernible differences in investor responses to top-mover stocks with market cap above or below \$300 million cutoff. Our assumption is that any differences in investor responses on either side of the cutoff, after controlling for the effect of market cap, should be only due to the impact of the Robinhood interface on investors' trading behavior.

Specifically, we select top mover stocks with market cap within a small bandwidth (e.g., \$50 million) around the \$300 million cutoff for our regression discontinuity analysis. Given the Robinhood top mover list displays the top 20 stocks with market cap above \$300 million that have the largest absolute percentage price moves measured from the previous market close price, we consider our treatment group for day t as stocks with market cap $\in$ (\$300 million, \$350 million] that ranked top 20 by absolute day-t overnight returns ($|R_{Treatment,t}^{Overnight}|$) among all stocks with market cap above \$300 million. Our control group then includes stocks with market cap $\in$ [\$250 million, \$300 million] that have absolute day-t overnight returns ($|R_{Control,t}^{Overnight}|$) close to stocks in the treatment group.[30] We also vary the bandwidth choices at three other levels (i.e., \$75, \$100, and \$125 million) to show the robustness of our results.

We exploit a sharp regression discontinuity design. Intuitively, this estimation exploits the discontinuity in information display at the \$300 million market cap threshold and tests for discontinuities in investor buying behavior around this threshold. We estimate the following pooled, cross-sectional sharp RD specification:

$$\begin{aligned} userchg_{it} = \beta_0 + \beta_1 I_{mktcap_{it}>\$300m} + \sum_{n=1}^{N} \beta_2^n (mktcap_{it} - \$300m)^n \\ + \sum_{n=1}^{N} \beta_3^n (mktcap_{it} - \$300m)^n \times I_{mktcap_{it}>\$300m} + \varepsilon_{it} \end{aligned} \tag{5}$$

where $userchg_{it}$ is the net buying activity of Robinhood investors measured by the intraday user changes on day t, $mktcap_{it}$ is the market cap of stock i at the market open of day t, and $I_{mktcap_{it}>\$300m}$ is an

[28]According to Robinhood Web Disclosures (https://cdn.robinhood.com/disclosures/WebDisclosures.pdf), "Robinhood uses a proprietary algorithm to display stocks with a market cap of more than \$300 million that have largest price movements as measured from the previous market close price" for the top mover list.

[29] Ideally, one could consider an alternative regression discontinuity design that exploits the threshold between the 20th and 21st top mover stocks. However, as we do not observe the actual ranking of price swings, and the ranking is likely to change throughout the day, the noise potentially introduced by using the approximated rankings may be too large to warrant a clean identification. Therefore, we opted to exploit the market cap threshold instead.

[30] For each stock in the treatment group, we find a matched stock that has the closest absolute return distance with the treated stock among all stocks with market cap $\in$ [\$250 million, \$300 million] that satisfy $0.5 \leq |R_{Treatment,t}^{Overnight}|/|R_{Control,t}^{Overnight}| \leq 2$. With the matching, the distributions of the absolute overnight returns are similar between the treatment group (mean: 0.090, median: 0.073, std dev: 0.059) and the control group (mean: 0.078, median, 0.062, std dev: 0.054). We also implement the same sharp RD test on the absolute overnight returns in Table A3 in the Internet Appendix, and there is not discontinuity in the absolute overnight returns at the \$300 million threshold.

indicator that equals one for stocks that have market cap at the market open of day t greater than $300 million. As controls, we include different polynomial functions of market cap ($N = 1,2,3$) so that the point estimate on the above-cutoff indicator variable (β_1) is identified under the assumption that the way that investor trading behavior is associated with market cap is not discontinuous exactly at the $300 million cutoff threshold for reasons besides the app interface.

The regression results in Table 7 show a discontinuous increase in Robinhood investor's buying activities when market cap surpasses the $300 million cutoff. The coefficient estimate (β_1) is positive and statistically significant. The results are robust with different bandwidth choices and with controls for the linear, quadratic, and cubic functions of the market cap. The graphical evidence corresponding to the three specifications with the $50 million bandwidth is presented in Figure 5. As shown in the figure, the intraday user change exhibits a clear jump at the market cap cutoff of $300 million.

To verify that we are not obtaining spurious estimates of the effects of the information display using the regression discontinuity design, we conduct a placebo test exploiting alternative market cap cutoffs at $250 million. For the placebo test using the $250 million cutoff, we select top mover stocks within a $50 million bandwidth around the $250 million cutoff (i.e. market cap ∈ ($200 million, $300 million]) and conduct the same regression estimation. Since none of the stocks in this placebo exercise would be displayed on the top mover list by Robinhood, the coefficient estimate of the treatment effects (β_1) should be zero. As shown in the regression results in the Internet Appendix Table A4, the coefficient estimate (β_1) is indeed statistically insignificant for $250 million cutoff.

Overall, these results are consistent with the idea that the Robinhood App's design impacts the trading decisions of its users.

3 Return Results

We find that relative to general retail investors, Robinhood users have more concentrated buying and selling. Concentrated buying is likely attention-driven and influenced by information display on the Robinhood interface. On days of extreme buying (i.e., herding events), Robinhood users could create price pressure (see Coval & Stafford, 2007). In this section, we examine the return patterns around such herding events.

3.1 Event Time Results

Our first analysis examines abnormal returns from herding event day -10 to event day 20. Day 0 is the herding day. Abnormal returns are calculated as the stock's return less the value weighted CRSP index. In Table 8, we report the mean abnormal return for each day and the buy-and-hold abnormal returns separately before and after the event. For example, the pre-event buy-and-hold abnormal returns are calculated as follows:

$$BHAR_{i\tau} = \prod_{t=\tau-10}^{\tau}(1 + R_{it}) - \prod_{t=\tau-10}^{\tau}(1 + R_{mt})\ . \tag{6}$$

We also report the percent of returns that are positive.

Standard errors are computed clustering on event day since we may have multiple events on the same day. Thus, the statistics underlying the mean daily returns lean on the reasonable assumption that returns are serially independent. The longer horizon abnormal returns are also clustered by event day, which partially corrects for the cross-sectional dependence issues. However, the standard errors are likely too small because of the overlapping nature of the returns at longer horizons. We address this econometric concern in the next section with a calendar-time trading strategy.

The buy-and-hold abnormal returns at longer horizons have the advantage that they accurately represent the return earned by investors. Cumulative abnormal returns (the sum of daily abnormal returns) are a positively biased representation of long-horizon abnormal returns in the presence of temporary price pressure effects or bid/ask bounce, both of which are likely issues in the stocks with the herding episodes we study.[31]

Prior to the herding event, stocks have abnormal returns near zero. Then a day or two before the herding event, average returns increase and become statistically significant. Next, the stocks experience an extremely positive return on the herding day – averaging 14%. Interestingly, many stocks have negative returns the day prior to and on the day of the herding event. This is consistent with our prior results documenting that extreme negative returns draw the attention of Robinhood users as well.

The pattern after the herding events is starkly different. Immediately after the herding event, returns turn significantly negative. After just five days, stocks in the Top 0.5% (rh_herd) experience negative abnormal returns of -3.5%. By the end of the 20-day period, the return decline totals almost 5%. These results are economically and statistically significant. These results are not driven by just a few stocks as almost two-thirds of rh_herd stocks have negative cumulative returns by the end of the 20 days.

To visualize these return patterns, in Figure 6, we plot the buy-and-hold abnormal returns for our rh_herd events. Results over the entire period are reported in Panel A, and $BHAR$s starting at Day 1 are reported in Panel B. The pattern of returns around herding events is quite clear. Robinhood users are attracted by extreme return events. Their coordinated buying leads to price pressure and then a subsequent

[31] To see this, consider a stock that cycles between ask and bid prices of \$10 and \$11 across three days, starting at the \$10 ask. Assume the market return is zero. The daily returns on day 1 are 10% (1/10) and day 2 are -9.1% (-1/11). The daily abnormal returns are 10% and -9.1%. The cumulative abnormal return is 0.9% (10% - 9.1%). The buy-and-hold abnormal return is zero ($11/10 \times 10/11 - 1$). While this example uses bid and ask prices for simplicity, the same logic applies to any mechanism that generates negative serial dependence in returns (e.g., temporary price pressure effects or liquidity provision).

poor return performance. We also find little evidence of unusual return movement beyond 20 days (see Internet Appendix Figure A3).

To systematically analyze the relation between the herding intensity and price reversal, we analyze stocks with a minimum of 100 Robinhood users and identify different sets of herding episodes by varying the daily user change ratio from 1.1 to 8.5 (i.e., from a 10% to 750% increase in users holding the stock). In Figure 7, Panel A, we plot the number of herding episodes (y axis in log scale) against the user change ratio used to identify the episodes (x axis). At a user change ratio of 1.1, we observe over 20,000 herding episodes; at a user change ratio of 8.5, we observe 45 episodes. In Panel B, we show the mean abnormal return that we document following these herding events grows from a statistically significant -1.8% at a user change ratio of 1.1 (> 20,000 events or about 36 events per day) to an extremely large -19.6% when we require a user change ratio of 8.5 or more (45 events or about 1 event every 12 days). There is a clear relation between the magnitude of the buying intensity and the subsequent reversal.

It's possible the return magnitudes of Figure 7, Panel B, increase as we move to more extreme cutoffs of the user change ratio because the identified stocks become smaller rather than the effect of the herding event. To rule out this size-based explanation, we recreate the results of Figure 7, but throw out stocks with market caps greater than $1 billion. This has the effect of ensuring mean and median size is the same across the user change ratio cutoffs, but we still see a dramatic increase in the observed return effects (see Internet Appendix Figure A4).

3.2 Calendar-Time Trading Strategy

To address the cross-sectional dependence issue underlying event-time analyses and to investigate the returns earned on a trading strategy that follows the herding episodes, we construct a calendar-time portfolio that invests $1 in each herding episode stock at the close of the event day and holds the stock for five days (without rebalancing).

We estimate the daily portfolio abnormal return ($alpha$) by estimating a regression of the portfolio excess return on Fama-French five-factor model plus a momentum factor:

$$R_{pt} - R_{ft} = \alpha + \beta\left(R_{mt} - R_{ft}\right) + \sum_{k=1}^{K} c_k F_t^k + e_{pt} \tag{7}$$

Where R_{ft} is the daily risk free return, R_{mt} is the value-weighted market index, and F_t^k are the k=1,K factor returns related to size, value, investment, profitability, and momentum (taken from Ken French's online data library). For completeness, we include results with just the market excess return (i.e., the CAPM) and four-factor alphas using market, size, value, and momentum factors.

Results in Table 9 are broadly consistent with the event-time analysis of the prior section. Over the sample period, the calendar time portfolio earns an economically large daily alpha of -55 to -61 bps (columns (1) to (3)), which is in line with the five-day event-time market-adjusted return of -3.55%. We also find the portfolio alphas are more negative during the 2020 pandemic period, ranging from -78 to -94 bps per day. The Robinhood user changes for herding stocks are also more dramatic during the pandemic period with a mean (median) $userratio$ of 1.75 (1.47) pre-COVID (prior to March 13, 2020) and 2.56 (1.83) post-COVID. However, the return effects of herding events are not unique for the pandemic period, the negative alphas are also sizable during pre-COVID period.

3.3 Regression Results

To further test the return patterns that we explore in event time, we regress daily stock returns for stock i on day t (R_{it}) on lags of the key herding variable (rh_herd) and controls:

$$R_{it} = a + \sum_{k=1}^{5} b_k \, rh\_herd_{i,t-k} + \sum_{k=1}^{5} c_k \, taq\_retimb_{i,t-k} + \sum_{k=1}^{5} d_k \, R_{i,t-k} + \mu_t + e_{it} \, . \quad (8)$$

The $\{b_k\}$ coefficients estimate the impact of herding events on returns 1 to 5 days after the event. We step in the control variables to assess how they interact with the herding events that we analyze. Robust standard errors are estimated with clustering by day, which addresses the cross-sectional dependence issue that effects the event-time graphs of the prior section.

The control variables include lagged retail order imbalance (taq_retimb) which has been shown to positively predict returns at short horizons in the US (Barber, Odean, & Zhu, 2008; Kaniel, Saar, & Titman, 2008; Kelley & Tetlock, 2013; Boehmer et al., 2020). The retail imbalance variable allows us to assess whether the poor returns we document are a manifestation of general retail order imbalance predicting returns during our sample period.

We also include lagged returns ($R_{i,t-k}$) to control for the well-documented tendency for return reversals at short horizons up to one month (French & Roll, 1986; Jegadeesh, 1990; Lehmann, 1990; Lo & Mackinlay, 1988, 1990; Campbell, Grossman & Wang, 1993) and that these effects are stronger in volatile markets including the 2020 pandemic period (Nagel, 2012; Drechsler, Moreira and Savov, 2020). These studies speculate that the origins of return reversals may emanate from overreaction, the rewards to providing liquidity provision, and/or more banal microstructure issues (e.g., prices bouncing between bid and ask prices). Attention-motivated trading leads to excessive buying and return reversals following attention-grabbing price increases, so may be a partial explanation for the returns associated with short-term contrarian strategies. Thus, the inclusion of lagged returns may be overcontrolling as both return reversals and the negative returns we document may have similar origins in attention-motivated trading.

Table 10 presents results for our herding measure based on the top 0.5% of daily Robinhood user changes. We present results for the full sample of herding events in columns (1)-(3). The last row of the table presents the summed coefficients on the herding indicator variables, which can be interpreted as the five-day abnormal return after the event. These five-day return estimates range from losses of 2.6% to 2.9%.

In Internet Appendix Table A5 we use more fine-grained controls for extreme positive and negative return moves and find similar five-day abnormal return estimates derived from the lags of the key rh_herd variable.

We separately analyze herding events that occur following a positive or negative overnight return using this regression model.[32] To analyze herding events that follow a positive overnight return, the key indicator variable takes on a value of one only if the overnight return (measured from close on day t-1 to open on day t) is positive and the change in users from the close on day t-1 to the close on day t is in the top 0.5% of stocks with positive user changes. The results are presented in columns (4) to (6). The herding events that are preceded by positive overnight returns predict somewhat stronger negative abnormal returns, ranging from -3.029% to -3.896%. We estimate a similar regression conditioning on negative overnight returns in columns (7) to (9). Though smaller in magnitude, we continue to observe negative abnormal returns for those herding events preceded by negative overnight returns with five-day abnormal returns ranging from -1.490% to -2.341%.

These analyses provide evidence that supports the conclusion that attention-motivated trading generates predictable poor returns.

3.4. TAQ Herding Events v. Robinhood Herding Events

In the prior sections, we develop a herding measure based on Robinhood user changes. In this section, we show that this herding measure is closely related to the herding measure developed by Barber, Lin, and Odean (BLO 2021) using TAQ retail trades. BLO document that stocks in the top quintile of retail standardized order imbalance and top quintile of abnormal retail volume earn dismal returns over the 2000-2019 sample period.

We modify the BLO measure to create a sample of Robinhood herding events and TAQ herding events for each day in the Robinhood sample period we analyze in this paper as follows. On each day, we use the standardized order imbalance (*SOI*) measure of BLO to construct quintiles and identify the stocks in the top quintile of retail order imbalance. Among stocks in this top quintile, we rank stocks based on abnormal retail volume (defined as retail volume on day t divided by mean volume from $t-20$ to $t-1$). We then create an indicator variable (taq_herd) that takes a value of one for the N stocks with the greatest abnormal retail volume with N equal to the number of Robinhood herding events that we identify on day t. Thus, for each day during our sample period we have an equal number of Robinhood herding events and TAQ herding events.

[32] We condition on overnight returns rather than close-to-close returns because overnight returns generally precede Robinhood user changes, which occur most commonly at or after the open of trading. If we condition on close-to-close returns, the herding event might cause the positive returns. When we condition on positive (negative) close-to-close returns, the five-day abnormal return estimated in column (4) and column (6) are -3.38% and -1.89%, respectively (both with $p<.01$).

By construction, we have an equal number of TAQ and Robinhood herding events. More importantly, 73% (3,567 of 4,884) of the stock-day observations are identical; there is substantial agreement in the herding events identified using Robinhood user changes and the BLO method using TAQ retail trades.

To explore the ability of the two measures to predict future returns, we modify the regression framework of the prior section to include both herding measures. In Table 11, column (1) replicates the main results using the rh_herd variable for easy comparison; column (2) uses the identical specification but replaces rh_herd with taq_herd. The taq_herd variable also predicts negative abnormal returns, but the magnitudes are smaller (-2.208% versus -2.942%). (See Internet Appendix Table A6, A7, and Figure A5 for event time and calendar time analyses of TAQ herding events.)

In column (3), we add an indicator variable that takes a value of one if TAQ retail order imbalance is positive (*taqpos*) and interact it with *rh_herd*. This analysis shows that the Robinhood herding events reliably predict five-day abnormal returns of -1.506% even when there is no net buying by retail investors in TAQ. The interaction effects (TAQ buying and Robinhood herding) are large and statistically significant, which suggests the return effects of herding events are larger when the general population of retail investors are also buying. In column (4), we find similar results when we control for lagged returns and order imbalance.

In column (5), we include rh_herd, taq_herd, and their interactions. These results indicate the rh_herd variable generally has larger predictive ability (-2.397% versus -1.213% for five-day abnormal returns). However, the interaction effects for the two herding variables are not significant which indicates the negative returns for events identified by both measures have particularly dismal returns. Recall that 73% of the events are common, so the stocks with both a Robinhood and TAQ herding event constitute a substantial proportion of the overall herding sample.

In summary, the Robinhood herding measure we identify is closely related to the BLO (2021) TAQ herding measure. The Robinhood herding measure has stronger ability to predict short-term negative returns than the TAQ measure. Both measures combine provide the strongest signal that returns will be negative in the coming days.

3.5 Subsample Analyses

To test the robustness of these patterns, we conduct a battery of robustness tests in Table 12. Because Robinhood users can trade with limited capital, low priced stocks are appealing to them, and we thus include them in our analysis. Might our results be driven by a tendency to observe closing prices at bid prices on the day of the herding event and thus, on average, negative returns the next day? The fact that most of the losses are observed during the next day (rather than overnight) suggests bid-ask bounce is *not* the main driver of the results. To further address this issue, we re-estimate the results using returns based on quote midpoints and find qualitatively similar results (Panel B). We also find reliably negative returns for stocks with prices in excess of $5, but the return magnitudes are smaller (Panel C).

We anticipate the negative returns we document will be present in small cap stocks but muted or nonexistent in large cap stocks, where retail trading is less likely to influence pricing. Consistent with this idea, we find stocks with market caps less than $1 billion generate larger abnormal returns (-3.8 to -4.3%) and larger stocks with more than $1 billion in market cap have no discernable return pattern (Panels D and E). Our main results include both stocks and other securities (e.g., ADRs and ETFs). We observe similar return patterns for these other securities only (Panel F).

We note that there was a large increase in both retail trading and Robinhood user holdings during the pandemic period (after March 13, 2020). Thus, we anticipate that the magnitude of the attention-induced subsequent poor performance will increase during this period, which is precisely what we observe (Panels G and H).

3.6 Sales Herding

As noted previously, attention asymmetrically affects buying and selling because investors can buy any stock but tend to sell only that which they own. Thus, driven by this theory of attention, our primary analysis focuses on the buying behavior of Robinhood investors. It's natural to wonder whether there are detectable return effects when we analyze sales herding events. To address this issue, we construct a sales herding variable that is analogous to our purchase herding variable. Specifically, we identify the securities in the bottom 0.5% of negative user changes as a percent of prior day user count to construct the indicator variable $rh\_negherd$.

It's noteworthy that these sales herding events have user changes that are much less dramatic than the purchase herding events. For the nearly 4,900 sales herding events, the mean $userratio$ is 0.87 and the median is 0.90 (see Internet Appendix Table A8). Consistent with the attention model, the sales herding is less dramatic than the purchase herding. It's also noteworthy that the sales herding events tend to follow the purchase herding events and are more clustered than purchase herding events. For example, we estimate a linear probability that regresses the sales herding variable ($rh\_negherd$) on lags of itself and lags of the purchase herding variable (rh_herd). The coefficients on the lagged rh_herd variable are all significant and economically large; at one- and two-day lags, the coefficients are 0.192 and 0.132 (see Internet Appendix Table A9), orders of magnitude larger than the baseline probability of 0.0012. Roughly half of the sales herding events are preceded by a purchase herding event in the prior five days.

We then analyze the returns on the sales herding stocks using the regression format of Table 10. We find evidence that the sales herding events have relatively modest negative abnormal returns of about 55 to 63 bps over the five days following the decrease in users (see Internet Appendix Table A10), and most of the negative abnormal returns follow sales herding events that occur after negative overnight returns (columns (6) to (9)). The return effects following sales herding events suggest that sales herding events are not likely to create price pressure as purchase herding events, but rather accelerate the poor performance that follow purchase herding events.

3.7 Aggregate Investor Experience: Average Purchase and Sales Price

While we have found that herding events predict negative returns going forward, it could be that the Robinhood community collectively still profits around these episodes. Enough investors could purchase the stock before the herding event for those users' profits to exceed any losses by later purchasers. To analyze this issue, we compare the average purchase and sales prices for all users. Specifically, we compute the purchase prices of Robinhood users during the event period, $\tau = -10, +20$. Define $u_{i\tau} - u_{i,\tau-1} = \Delta u_{i\tau}$ as the change in users for stock *i* on event day τ. For event *j*, we calculate the average purchase price for days on which we observe users increases ($\Delta u_{it} > 0$) as:

$$PPrc_j = \frac{u_{i,-11}P_{i,-11} + \sum_{\tau=-10}^{+20} \mathrm{UI}_{i\tau}\Delta u_{i\tau}P_{i\tau}}{u_{i,-11} + \sum_{\tau=-10}^{+20} \mathrm{UI}_{i\tau}\Delta u_{i\tau}}. \tag{9}$$

The indicator $UI_{i\tau}$ equals one on days when the change in users is positive, $\Delta u_{i\tau} > 0$. Note that all users owning stock *i* at the close on day -11, before the event period begins, are assumed to have purchased at the closing price on day -11, i.e., $P_{i,-11}$.

We then compute the analogous calculation for average sales price on days when we observe a decrease in users. For any shares that are not sold during the event period, we assume they are sold at the end of the event window:

$$SPrc_j = \frac{u_{i,20}P_{i,20} + \sum_{\tau=-10}^{+19}(1 - \mathrm{UI}_{i\tau})\Delta u_{i\tau}P_{i\tau}}{u_{i,20} + \sum_{\tau=-10}^{+19}(1 - \mathrm{UI}_{i\tau})\Delta u_{i\tau}}. \tag{10}$$

The profitability of event j is then calculated as the ratio of the sales and purchase price:

$$PrcRatio_j = \frac{SPrc_j}{PPrc_j} - 1 \tag{11}$$

The $PrcRatio_j$ represents the returns earned by the Robinhood community on event j.

In addition to computing the raw return, we also compute returns that adjust for market movements. For each day during the event, we consider a counterfactual in which the Robinhood community buys or sells the equivalent amount of capital in an S&P 500 ETF, specifically SPY. This provides us with a market price ratio ($MktPrcRatio_j$) for each stock-event, which we can use to benchmark the price ratio for the event stock. We report results in Figure 8.

On average, using raw returns, we find that the Robinhood community loses approximately 4.3% during each herding event. After adjusting for the market return, losses are about 5.5%. Both results are highly significant both statistically and economically. There are slight differences between the pre-COVID-19 and COVID-19 periods, although both suggest similar outcomes. Overall, these findings suggest that extreme herding causes negative wealth outcomes for the overall Robinhood community.[33]

4 Short Trading and Herding Events

In the prior section, we found that Robinhood users' herding can lead to price pressure on those stocks and subsequent poor performance. Our test setup uses data available before market closes, and therefore implies that the negative returns are tradable. In other words, other market participants could trade against Robinhood users' order flow in these herding situations in an attempt to profit from Robinhood users' correlated trading. In some sense, the disclosure of the stock holdings by Robinhood is similar to the disclosure of positions by mutual funds, hedge funds, and others through Form 13F. Hedge funds have argued that these disclosures enable other market participants to profit from their private information and to front-run their positions.[34] Researchers have documented that these disclosures are used by other market participants (e.g. Brown & Schwarz, 2020). Thus, it is somewhat surprising that Robinhood would voluntarily make holdings data available to other market participants who might profit from the data at the expense of Robinhood users.[35]

In this section, we evaluate whether other market participants do attempt to profit from the predictably negative returns that follow Robinhood herding events. To do so, we collect data from FINRA which provides daily short trade data as noted by Boehmer and Song (2020). We compute the abnormal short volume on herding day t as the ratio of the number of short trades on day t divided by the average number of short trades from days $t-20$ to $t-1$. We also winsorize the top 0.1% to control for extreme outliers.

To begin our analysis, we examine the relation between top stock performance, Robinhood herding, and short trading. Specifically, we calculate the daily return rank for each stock since we expect abnormal short trading to be correlated with high returns. We then flag whether the stock was also herded into by Robinhood users any day during the period using our 0.5% cutoff (rh_herd). We then calculate the average change in abnormal short volume for each return rank for the non-Robinhood herding and Robinhood

[33] In Internet Appendix Figure A6, we show that in the average event, about 60% of investors who buy during these herding episodes buy at a price that is higher than the observed price 20 days after the herding event. Sellers fare better, but buyers outnumber sellers by approximately 18.7 million to 3.5 million.

[34] It is not clear whether hedge fund 13Fs contain any private information. Griffin and Xu (2009) find that 13F disclosures have no alpha as of their effective data. Brown and Schwarz (2020) find no alpha as of the date filed with SEC EDGAR. Aiken, Clifford, & Ellis (2013) find that, after controlling for biases, hedge fund returns themselves have little alpha.

[35] Robinhood ceased releasing user data in August 2020, stating that the way the data is sometimes reported by third parties "could be misconstrued or misunderstood" and does not represent the company's user base. (See https://www.foxbusiness.com/markets/robinhood-to-stop-sharing-of-apps-popular-stocks.)

herding groups separately. In Figure 9, we plot the average values for ranks 1-5, 6-10, 11-15, 16-20, 21-25, 26-50, and 51-100.

We find two clear patterns. First, abnormal short volume are indeed higher for stocks that had a high daily return rank. Second, we see in all cases abnormal short volume for stocks herded into by Robinhood investors is far higher than those stocks not herded into Robinhood investors. This is consistent with market participants knowing the pattern of Robinhood return reverses and, more broadly, consistent with market participants using the disclosed Robinhood position information.

We then examine short-interest changes in a multivariate setting. Each period, we regress abnormal short trading on either the daily excess return or the return ranking (ret_rank). We also include abnormal news coverage to control for attention that news may have brought and caused. Our primary variables of interest are either the average user change ratio during the period ($rh\_chgratio$) or herding of the stock by Robinhood users (rh_herd). We average across periods and obtain standard errors and t-values via Fama-Macbeth (1973).

In Table 13, we find results consistent with the previously reported figure. A 100% increase in (doubling) the average change ratio leads to at least a 637% increase in short trading (Column (1)), which is both statistically and economically significant. If a stock is herded by Robinhood users (Columns (2)), abnormal short volume is approximately 855% higher than if the stock is not herded. Overall, these results again suggest strongly that market participants examined Robinhood ownership data, knew about the subsequent poor performance caused by Robinhood herding, and traded against Robinhood order flow.

5 Conclusion

The stated mission of the Robinhood brokerage is to "democratize finance for all…[and] make investing friendly, approachable, and understandable." Robinhood facilitates "friendly, approachable, and understandable" investing by offering a simple downloadable app that makes trading incredibly easy. The app displays only a small fraction of the stock level indicators that other brokerage platforms provide. Instead, the app highlights easily understood lists of stocks such as the "Top Mover" list of stocks with the largest price moves on the current day.

We argue that the combination of simplified information display and inexperience exacerbate attention-driven buying by Robinhood users. Heightened attention-driven buying leads to more concentrated trading by Robinhood users than other retail investors and contributes to buy-side herding events that are usually followed by negative returns. For example, the top 0.5% of stocks bought by Robinhood users each day experiences negative average returns of approximately 5% over the next month. More extreme herding events are followed by negative average returns of almost 20%.

Robinhood has been successful in its stated mission in as much as it has attracted 13 million users (as of 2020). Half of these are first-time investors, who are likely in the long run to benefit from participating in markets. Robinhood attracts investors by reducing frictions and promoting simplicity. While a lack of frictions encourages market participation, it also makes speculative trading easy, which can lead to lower investment returns (Odean, 1999; Barber and Odean, 2000; Barber, Lee, Liu, and Odean, 2009). Even in an industry that uses complexity to obscure risks and costs (Carlin, 2009; Henderson and Pearson, 2011; Célérier and Vallée, 2017; Gao et. al., 2021), simplicity is not problem free. As we show, simply focusing the attention of many investors on a small number of stocks can promote herding behavior that affects market returns and redounds to the investors' detriment. Thus, while it is important that investors have access to transparent, pertinent information, disclosure alone is not sufficient to assure good investor outcomes; how information is displayed influences decisions in ways that can both help and hurt investors.

References

Aiken, A., Clifford, C., & Ellis, J. (2013). Out of the Dark: Hedge Fund Reporting Biases and Commercial Databases. *Review of Financial Studies*, *26*, 208–243.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2009). Just How Much do Individual Investors Lose by Trading?. *Review of Financial Studies, 22*(2), 609-632.

Barber, B. M. and Lin, S. and Odean, T. (2021). Resolving a Paradox: Retail Trades Positively Predict Returns but are Not Profitable. *Working Paper.*

Barber, B. M., & Loeffler, D. (1993). The "Dartboard" Column: Second-Hand Information and Price Pressure. *Journal of Financial and Quantitative Analysis*, 273-284.

Barber, B. M., & Odean, T. (2000). Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors. *Journal of Finance, 55*(2), 773-806.

Barber, B. M., & Odean, T. (2002). Online Investors: Do the Slow Die First?. The Review of Financial Studies, 15(2), 455-488.

Barber, B. M., & Odean, T. (2008). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, *21*(2), 785-818.

Barber, B. M., Odean, T., & Zhu, N. (2008). Do Retail Trades Move Markets?. *Review of Financial Studies*, *22*(1), 151-186.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (Forthcoming). Tracking Retail Investor Activity. *Journal of Finance.*

Boehmer, E., & Song, W. (2020) Smart Retail Traders, Short Sellers, and Stock Returns. *Working Paper*.

Bolster, P., Trahan, E., & Venkateswaran, A. (2012). How Mad is Mad Money? Jim Cramer as a Stock Picker and Portfolio Manager. *Journal of Investing*, *21*(2), 27-39.

Brown, S., & Schwarz, C. (2020). Do Market Participants Care about Portfolio Disclosure? Evidence from Hedge Funds' 13F Filings. *Working Paper*.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading Volume and Serial Correlation in Stock Returns. *Quarterly Journal of Economics*, *108*(4), 905-939.

Carlin, B. I. (2009). Strategic Price Complexity in Retail Financial Markets. *Journal of Financial Economics*, *91*(3), 278-287.

Célérier, C., & Vallée, B. (2017). Catering to Investors Through Security Design: Headline Rate and Complexity. *Quarterly Journal of Economics, 132*(3), 1469-1508.

Cheng, P., Murphy, T. J., & Kolanovic, M. (2020). Following the Robinhood Money. *JP Morgan report*.

Choi, J. J., Laibson, D., & Madrian, B. C. (2010). Why does the Law of One Price Fail? An Experiment on Index Mutual Funds. *Review of Financial Studies*, *23*(4), 1405-1432.

Cookson, J. A., Engelberg, J., & Mullins, W. (2021). Echo Chambers. *Working Paper*.

Coval, J., & Stafford, E. (2007). Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics, 86*(2), 479-512.

Cronqvist, H., & Thaler, R. H. (2004). Design choices in privatized social-security systems: Learning from the Swedish experience. *American Economic Review*, *94*(2), 424-428.

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. Journal of Finance, 66(5), 1461-1499.

Da, Z., Hua, J., Hung, C., & Peng, L. (2020). Market Returns and a Tale of Two Types of Attention. *Working Paper*.

Da, Z., Larrain, B., Sialm, C., & Tessada, J. (2018). Destabilizing Financial Advice: Evidence from Pension Fund Reallocations. *Review of Financial Studies*, *31*(10), 3720-3755.

Drechsler, I., Moreira, A., & Savov, A. (2020). Liquidity and Volatility. *Working Paper*.

Eaton, G., Green, C., Roseman, B., & Wu, Y. (2021). Zero-Commission Individual Investors, High Frequency Traders, and Stock Market Quality, *Working paper*.

Engelberg, J., Sasseville, C., & Williams, J. (2012). Market Madness? The Case of Mad Money. *Management Science*, *58*(2), 351-364.

Fama, E., & MacBeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy, 81*, 607–636.

Fedyk, A. (2019). Front page news: The effect of news positioning on financial markets. *Working paper*.

French, K. R., & Roll, R. (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics, 17*(1), 5-26.

Frydman, C., & Wang, B. (2020). The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment. *Journal of Finance*, *75*(1), 229-276.

Gao, P., Hu, A., Kelly, P., & Peng, C. (2021). Exploited by Complexity. *Working Paper*.

Greenwood, R., & Nagel, S. (2009). Inexperienced Investors and Bubbles. *Journal of Financial Economics*, *93*(2), 239-258.

Griffin, J., & Xu, J. (2009). How Smart are the Smart Guys? A Unique View from Hedge Fund Stock Holdings. *Review of Financial Studies, 22*, 2331-2370.

Henderson, B. J., & Pearson, N. D. (2011). The Dark Side of Financial Innovation: A Case Study of the Pricing of a Retail Financial Product. *Journal of Financial Economics, 100*(2), 227-247.

Hong, C. Y., Lu, X., & Pan, J. (2019). The Economic Impact of Distributing Financial Products on Third-Party Online Platforms. *Working Paper.*

Jegadeesh, N. (1990). Evidence of Predictable Behavior of Security Returns. *Journal of Finance*, *45*(3), 881-898.

Jegadeesh, N., & Titman, S. (1995). Overreaction, Delayed Reaction, and Contrarian Profits. *Review of Financial Studies*, *8*(4), 973-993.

Kaniel, R., & Parham, R. (2017). WSJ Category Kings–The Impact of Media Attention on Consumer and Mutual Fund Investment Decisions. *Journal of Financial Economics*, *123*(2), 337-356.

Kaniel, R., Saar, G., & Titman, S. (2008). Individual Investor Trading and Stock Returns. *Journal of Finance*, *63*(1), 273-310.

Kahneman, D. (1973). *Attention and Effort*. NJ: Prentice-Hall.

Kahneman, D. (2011). *Thinking, Fast and Slow.* NY: Farrar, Straus and Giroux.

Keasler, T. R., & McNeil, C. R. (2010). Mad Money Stock Recommendations: Market Reaction and Performance. *Journal of Economics and Finance, 34*(1), 1-22.

Kelley, E. K., & Tetlock, P. C. (2013). How Wise are Crowds? Insights from Retail Orders and Stock Returns. *Journal of Finance*, *68*(3), 1229-1265.

Kronlund, M., Pool, V. K., Sialm, C., & Stefanescu, I. (2020). Out of Sight No More? The Effect of Fee Disclosures on 401(k) Investment Allocations. *Working Paper*.

Liang, B. (1999). Price Pressure: Evidence from the "Dartboard" Column. *Journal of Business*, *72*(1), 119-134.

Lehmann, B. N. (1990). Fads, Martingales, and Market Efficiency. *Quarterly Journal of Economics, 105*(1), 1-28.

Lo, A. W., & MacKinlay, A. C. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, *1*(1), 41-66.

Lo, A. W., & MacKinlay, A. C. (1990). When are Contrarian Profits Due to Stock Market Overreaction? *Review of Financial Studies*, *3*(2), 175-205.

Loos, B., Meyer, S., & Pagel, M. (2020). The Consumption Effects of The Disposition to Sell Winners and Hold Losers. *Working Paper*.

Nagel, S. (2012). Evaporating liquidity. *Review of Financial Studies, 25*(7), 2005-2039.

Niessner, M. (2015). Strategic Disclosure Timing and Insider Trading. *Working Paper.*

Odean, T. (1999). Do Investors Trade too Much?. *American Economic Review, 89*(5), 1279-1298.

Oprea, R. (Forthcoming). What Makes a Rule Complex? *American Economic Review*.

Ozik, G., Sadka, R., & Shen, S. (2020). Flattening the Illiquidity Curve: Retail Trading During the COVID-19 Lockdown. *Working Paper*.

Pedersen, L. H. (2021). Game on: Social networks and markets. *Working Paper*.

Seasholes, M. S., & Wu, G. (2007). Predictable Behavior, Profits, and Attention. *Journal of Empirical Finance*, *14*(5), 590-610.

So, E. C., & Wang, S. (2014). News-driven Return Reversals: Liquidity Provision Ahead of Earnings Announcements. *Journal of Financial Economics*, *114*(1), 20-35.

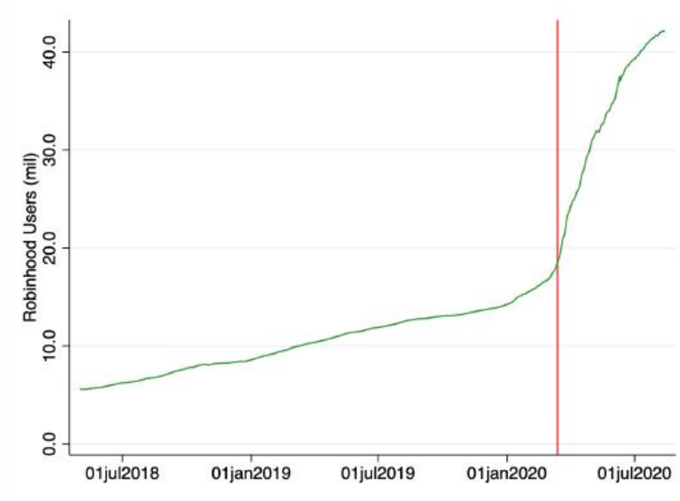
Tetlock, P. C. (2011). All the News That's Fit to Reprint: Do Investors React to Stale Information?. *Review of Financial Studies*, *24*(5), 1481-1512.

Umar, T. (2020). Complexity Aversion When Seeking Alpha. *Working Paper*.

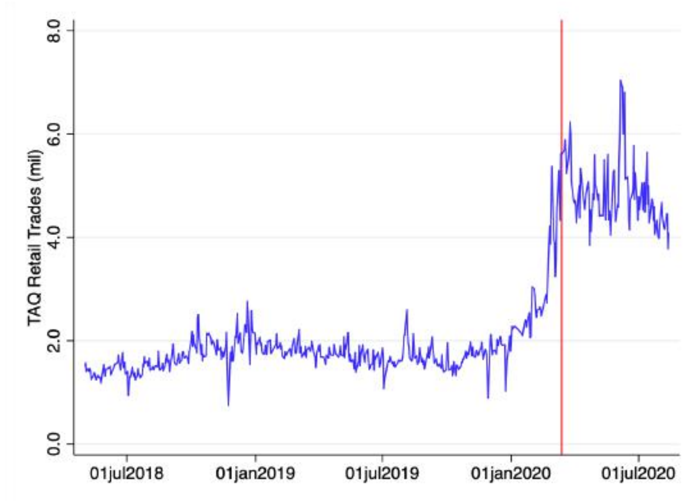
Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy, 110*(4), 825-853.

Welch, I. (2020). Retail Raw: Wisdom of the Robinhood Crowd and the COVID Crisis. *Working Paper*.

Panel A: Robinhood Total User Stock Holdings



Panel B: TAQ Total Retail Trades



Panel C: Summed Absolute Robinhood User Changes and TAQ Net Buying

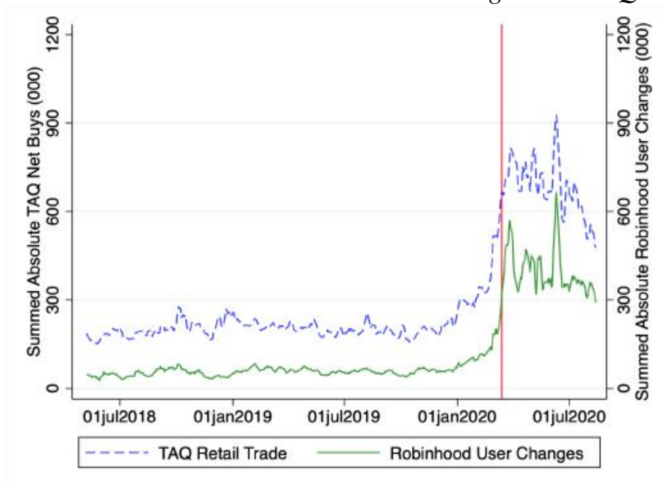


Figure 1: Robinhood and TAQ Retail Trading

Panel A plots the total number of Robinhood stock holdings. Panel B plots the total number of TAQ Retail Trades. Panel C plots the total number of absolute Robinhood user changes (green) and absolute TAQ net buys (blue) as a five-day moving average. The red line depicts the date when the COVID national emergency was declared in the US (March 13, 2020).

Panel A: Concentration of Buying *Panel B: Concentration of Selling*

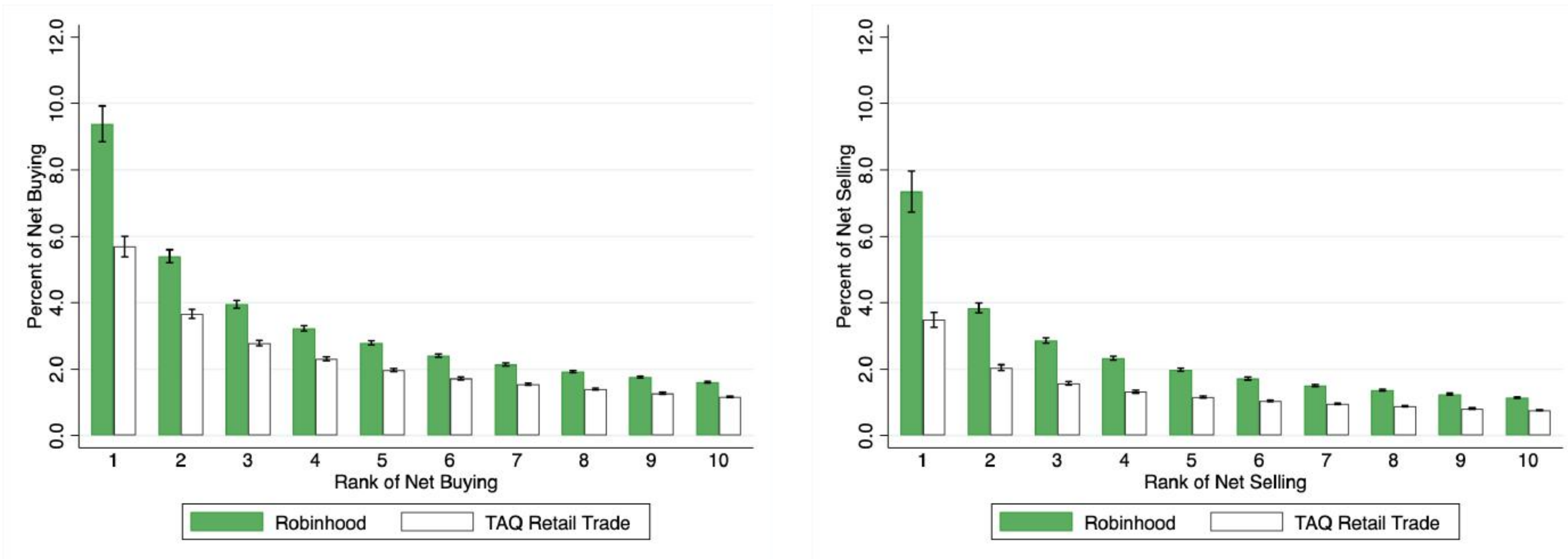
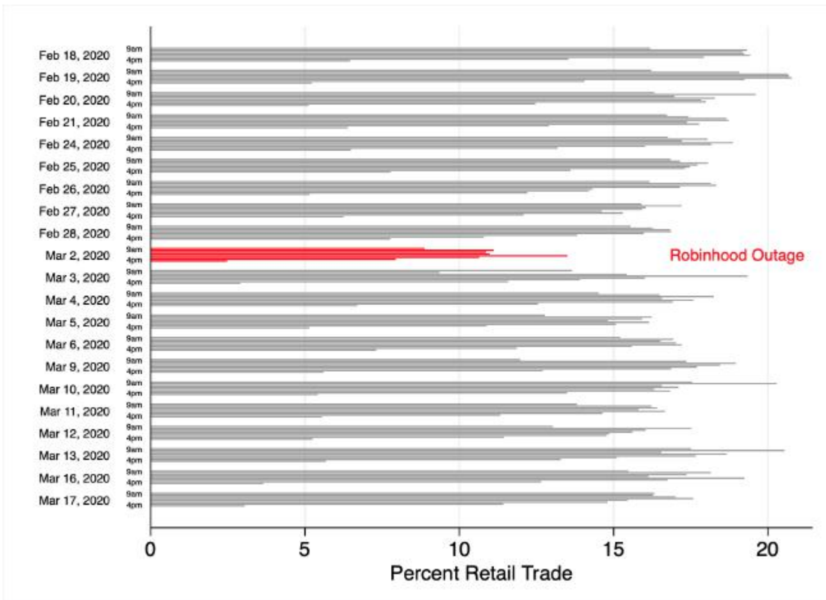


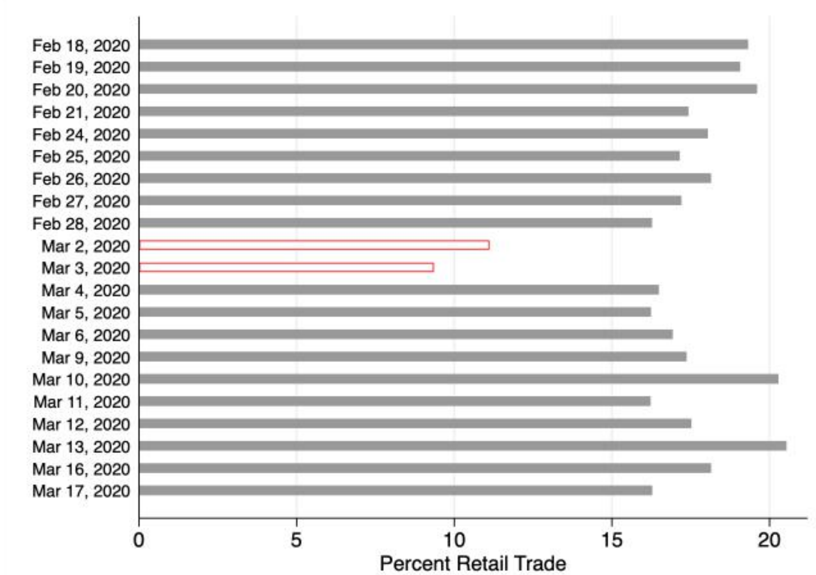
Figure 2: The Concentration of Buying and Selling

In Panel A, the figure depicts the mean daily percent of net buying that is observed in the stock with ranks from 1 to 10, with the rank of 1 being the stock with the most net buying. In Panel B, the figure depicts the mean daily percent of net selling that is observed in the stock with ranks from 1 to 10, with the rank of 1 being the stock with the most net selling. Whiskers depict 95% confidence intervals based on standard errors across days. In Robinhood, net buying is user changes. In TAQ, net buying is retail buys less retail sells.

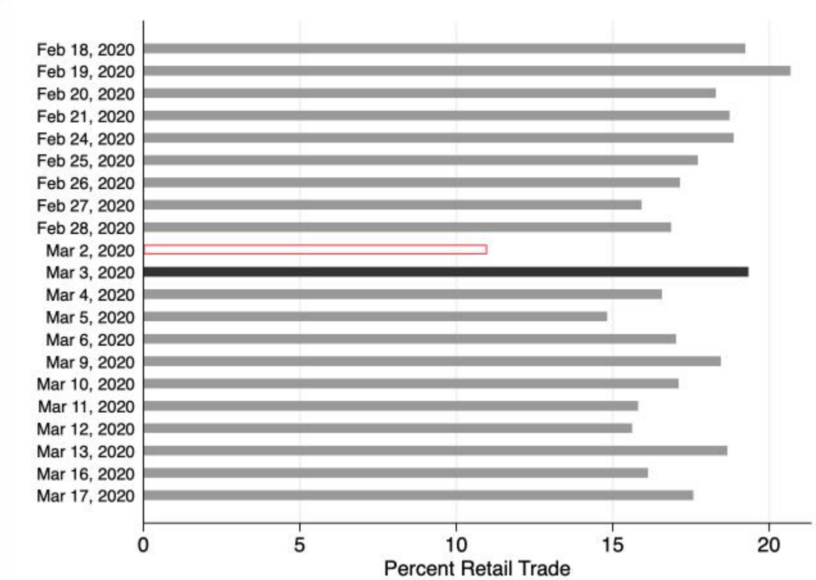
Panel A: Hourly Volume (March 2, 2020, Full Day Outage)



Panel B. 10-11 am Volume (March 3, 2020, 10 am Outage)



Panel C. Noon-1 pm Volume (March 2, 2020, Full Day Outage; March 3, 2020, Noon Repair)



Panel D. 11:35 am to 12:40 pm Volume (June 18, 2020, 11:35 am Outage)

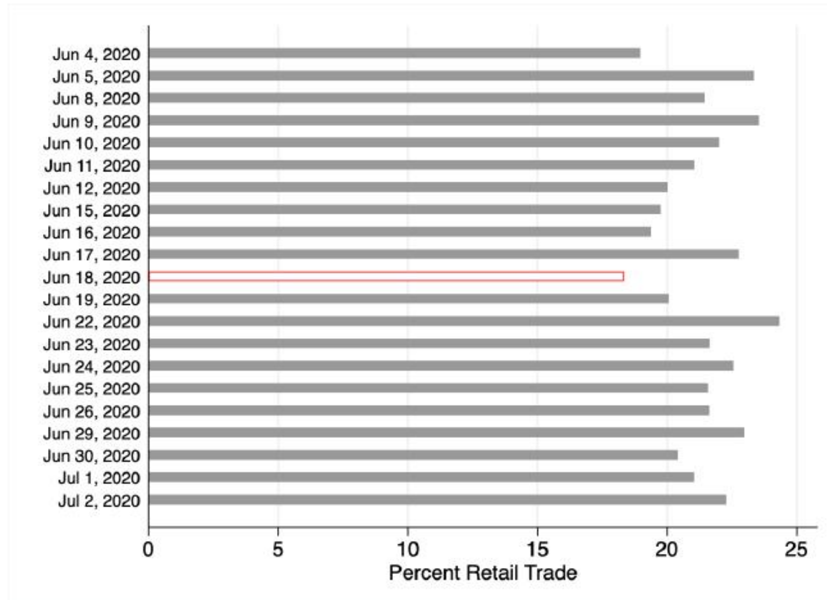
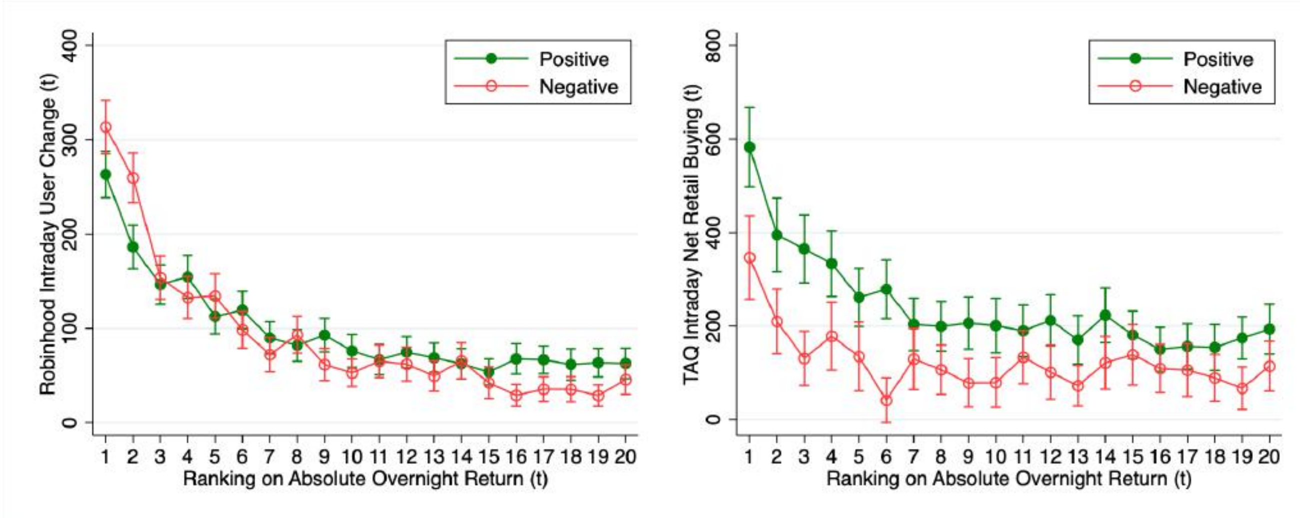


Figure 3: The Effect of Robinhood Outages on Retail Trade

In Panels A to C, the sample consists of the 50 most popular Robinhood stocks on February 28, 2020. In Panel A, March 2, 2020, is the day of a Robinhood outage (red bars). In Panel B, March 2, 2020, is the day of a Robinhood outage (white bar). A second shorter outage occurred around 10:00 am on March 3, 2020, and lasted for a bit more than an hour (white bar). In Panel C, March 2, 2020, is the day of a Robinhood outage (white bar). Robinhood tweeted all systems were fully restored at 11:55 am on March 3, 2020 (black bar). In Panel D, the sample consists of the 50 most popular Robinhood stocks as of June 17, 2020. The outage occurs between 11:30-12:30 on June 18, 2020 (white bar).

Panel A: Top Mover Rankings Sorted on Absolute Overnight Return



Panel B: Top Mover Rankings Sorted on Absolute Daily Return

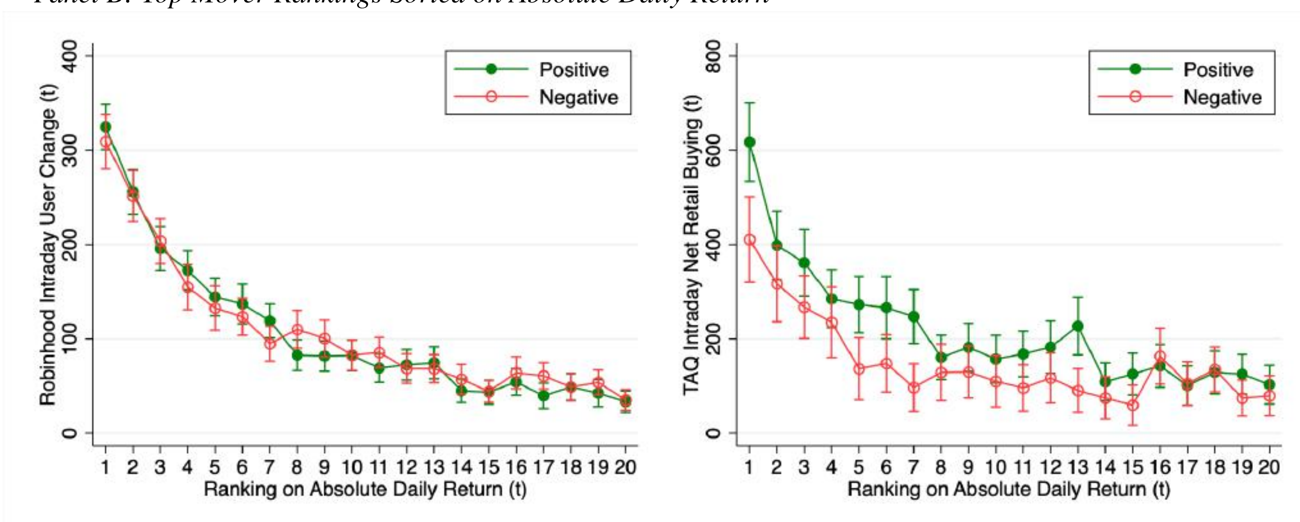


Figure 4: Mean Robinhood (RH) Intraday User Change vs. TAQ Intraday Net Retail Buying on Top Mover Rankings

The figures present the mean Robinhood intraday user change and TAQ intraday net retail buying of day t against top mover rankings based on different measures. The rankings are sorted for stocks with market cap above \$300 million at the market open of day t. Panel A sorts top movers on absolute overnight returns of day t; Panel B sorts top movers on absolute daily returns of day t. RH Intraday User Change measures the change in RH users from first time stamp that excludes market open trades to the last time stamp before market closes. TAQ Intraday Net Retail Buying is TAQ retail buying minus TAQ retail selling (with short trades removed following Boehmer and Song (2020)) during the market trading hours. The error bar represents the 90% confidence interval for the mean.

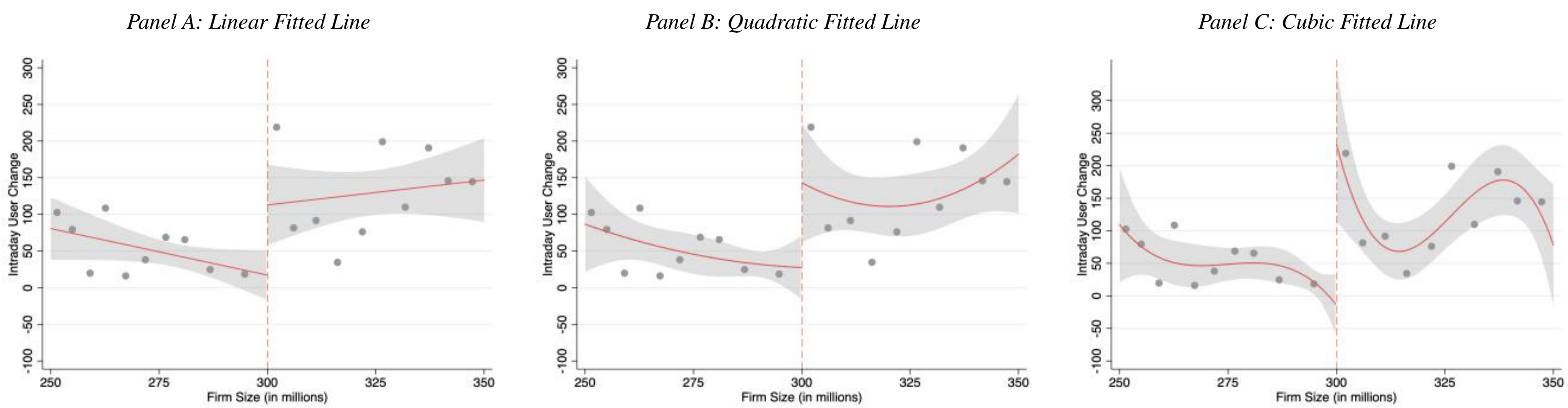


Figure 5: Robinhood (RH) Intraday User Change for Top Movers around the $300 million Market Cap Cutoff

This figure shows a binned scatter plot for the Robinhood intraday user change against market cap for stocks included in the regression discontinuity analysis as described in Table 7. The red lines are the linear (Panel A), quadratic (Panel B), and cubic fitted lines (Panel C) estimated separately for market cap above and below the $300 million cutoff. The shaded areas indicate the 90% confidence interval.

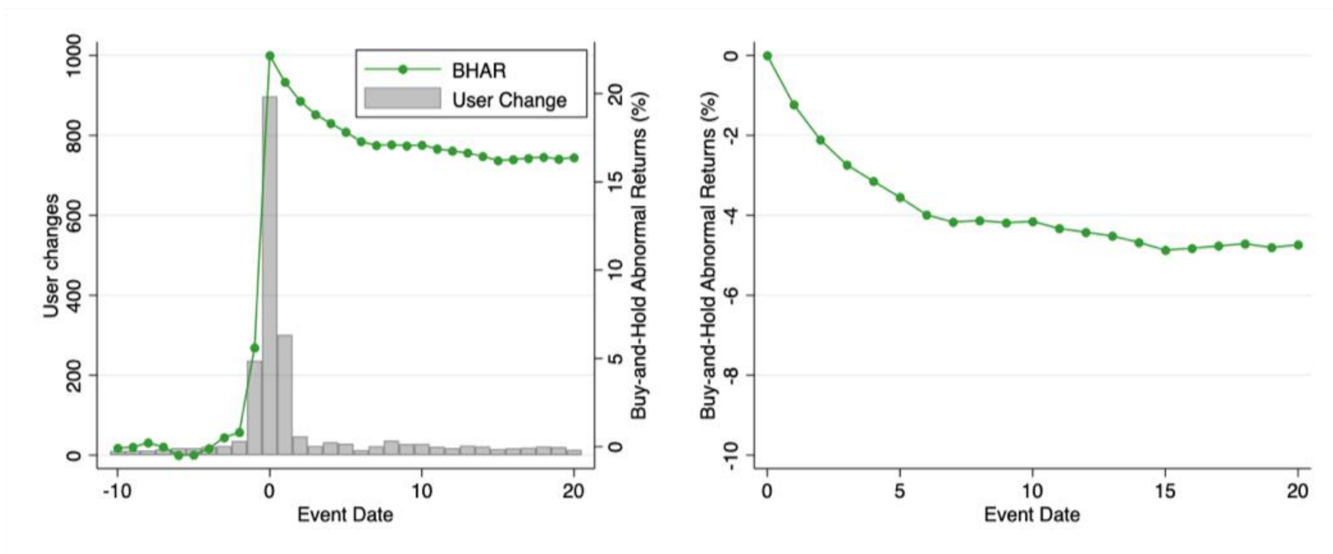
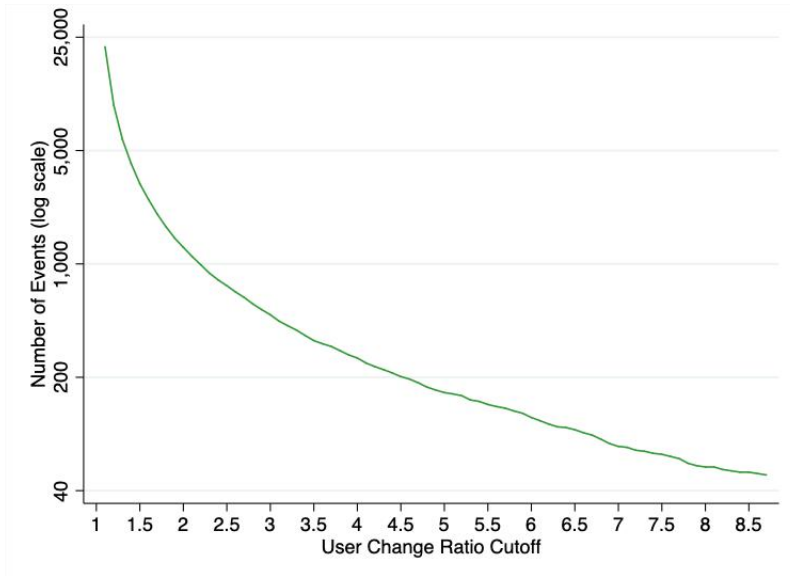


Figure 6: Returns around Herding Events

The figure on the left depicts the mean buy-and-hold abnormal returns (BHARs) and Robinhood user changes from ten days before to 21 days after herding events. The green line represents the BHAR, whereas the grey bars represent user changes. The figure on the right displays post-event mean BHARs starting from day 0. Herding events are defined as the top 0.5% of stocks with positive user change ratio on day 0 and a minimum of 100 users on the prior day.

Panel A: Herding Intensity and Daily User Change Ratio Cutoff



Panel B: Buy-and-hold Abnormal Returns (BHAR) and Daily User Change Ratio Cutoff

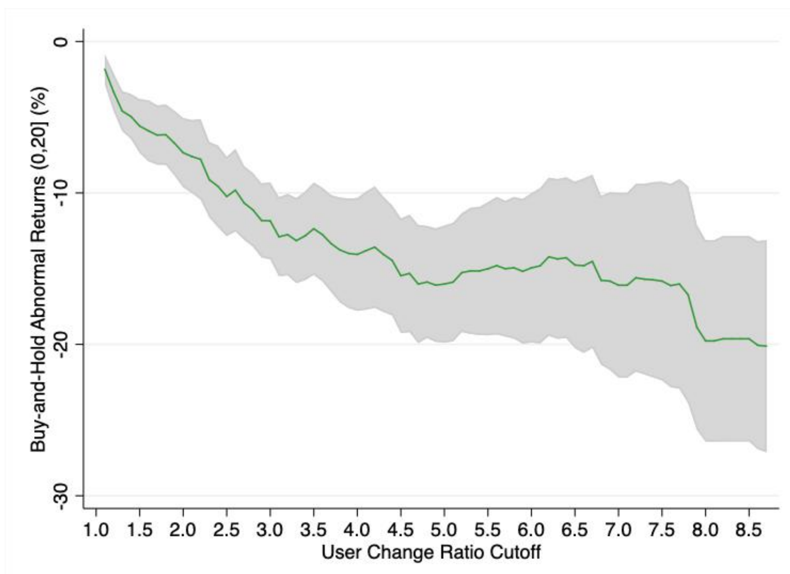


Figure 7: Herding Intensity and Price Reversal

The figures show how herding intensity and 20-day buy-and-hold abnormal returns vary with the daily user change ratio cutoff to identify the episodes for stocks with at least 100 Robinhood users. Panel A plots the number of herding episodes (in log scale) against the user change ratio cutoff. Panel B plots the 20-day buy-and-hold abnormal returns (%) against the user change ratio cutoff. The daily user change ratio cutoff varies from 1.1 to 8.5.

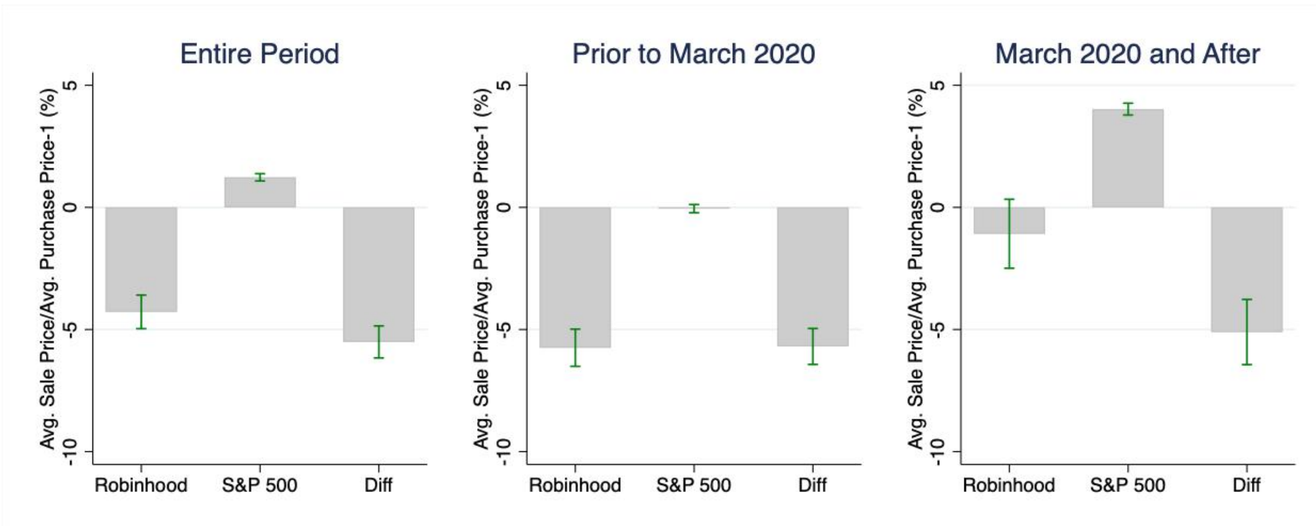


Figure 8: Aggregate Investor Experience: Average Profitability

The figure presents the average profitability of theRobinhood user community's actual trades across the herding episodes, their counterfactual trades on S&P 500 ETF during the herding episodes, and the difference between these two. For each herding episode, we compute the weighted average purchase (sales) price of Robinhood users during the event period [-10,20]. The profitability is calculated as (average sales price/average purchase price - 1). A positive profitability indicates that the Robinhood community profited from that herding episode. The counterfactual trades on S&P 500 ETF assume that the Robinhood community purchases or sells the equivalent amount of capital in an S&P 500 ETF. The error bar represents the 95% confidence interval for the average.

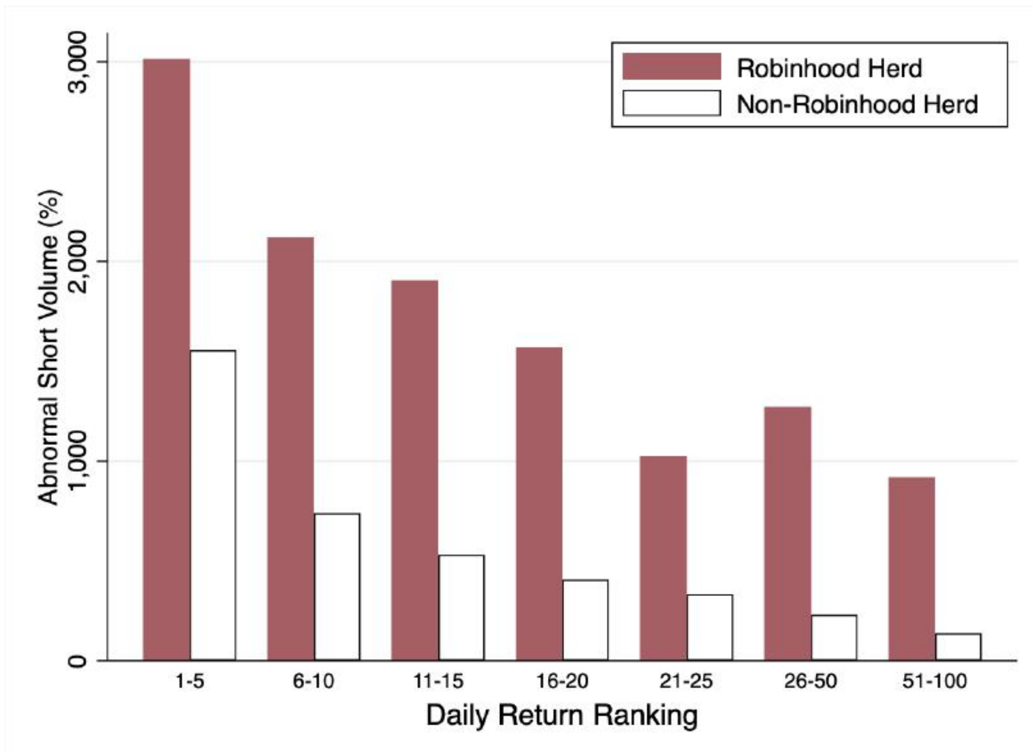


Figure 9: Short Interest Changes and Robinhood Trading

The figure depicts the average abnormal short trading (%). Short trading is measured using data from FINRA which provides daily short trade data as noted by Boehmer and Song (2020). We measure abnormal short volume on day t as the ratio of short trades on day t to the prior 20 day ($[t-20, t-1]$) average of the number of short trades. To control for outliers, we winsorize the top 0.1% of observations. The x-axis is daily return ranking of day t. Robinhood Herd are the top 0.5% of stocks with positive user change ratio on day t and a minimum of 100 users on day $t-1$.

Table 1: Summary Statistics

Panel A presents summary statistics across stock-day observations. Panel B sums variables by day and then averages across days. The variables are *users_close* (last observed user count for a stock prior to the 4 pm ET close), *users_last* (last user count of the day), *userchg* (daily change in *users_close*), *userratio* ($users\_close(t)/users\_close(t-1)$), *prc* (closing price), *size* (market cap in millions), *ret* (daily return), *openret* (overnight return), *dayret* (daytime return), *daily_buys* (number of TAQ daily retail buys), *daily_sells* (number of TAQ daily retail sells), *net_buys* ($daily\_buys - daily\_sells$), *taq_retimb* ($net\_buys/(daily\_buys + daily\_sells)$), and #*stocks* (number of stocks with users reported on Robintrack).

variable	N	mean	sd	min	p25	p50	p75	max
Panel A: Stock-Day Observations								
users_close	3,952,749	2,064.27	15,422.64	0.00	35.00	160.00	674.00	990,059.00
users_last	4,067,791	2,061.48	15,419.80	0.00	35.00	160.00	673.00	990,587.00
userchg	3,851,419	9.46	245.07	-19,643.00	-1.00	0.00	1.00	85,193.00
userratio	3,745,652	1.01	0.32	0.00	1.00	1.00	1.01	263.67
prc	3,765,043	52.47	1,828.84	0.04	10.38	23.71	46.81	344,970.00
size ($mil)	3,625,145	5,674.62	29,154.96	0.00	108.25	498.94	2,416.15	1,581,165.00
ret (%)	3,764,157	0.04	4.21	-91.79	-0.98	0.02	0.99	897.73
openret (%)	3,674,652	0.10	2.64	-88.60	-0.40	0.03	0.55	563.90
dayret (%)	3,696,846	-0.05	3.41	-87.59	-0.91	0.00	0.78	841.18
daily_buys	3,586,637	200.69	1,112.99	0.00	9.00	34.00	117.00	185,930.00
daily_sells	3,586,637	178.97	875.49	0.00	8.00	34.00	115.00	113,152.00
net_buys	3,586,637	21.72	367.93	-30,246.00	-7.00	0.00	10.00	86,640.00
taq_retimb	3,585,659	0.01	0.35	-1.00	-0.14	0.00	0.16	1.00
Panel B: Daily Observations (summed variable averaged across days)								
#*stocks*	549	7,211.01	741.42	5,805.00	6,559.00	7,199.00	8,054.00	8,131.00
users_close (mil.)	549	14.86	9.93	1.32	8.27	11.83	15.22	42.14
users_last (mil.)	549	14.89	9.93	5.58	8.27	11.83	15.23	42.16
userchg (000)	535	68.11	112.57	-48.83	14.70	23.74	54.11	810.40
daily_buys (000)	549	1,276.35	739.49	387.97	824.86	923.12	1,297.50	3,952.49
daily_sells (000)	549	1,137.28	585.45	360.04	778.65	872.51	1,181.31	3,174.90
net_buys (000)	549	139.07	171.91	-67.98	38.44	63.30	133.80	1,005.84

Table 2: Summary Statistics for Robinhood Herding Events

Herding events as securities in the top 0.5% of positive user change ratio on day t and a minimum of 100 users on day $t-1$. Summary statistics are presented for stock-days that meet these herding definitions. The variables are *users_close* (last observed user count for a stock prior to the 4 pm ET close), *users_last* (last user count of the day), *userchg* (daily change in *users_close*), *userratio* ($users\_close(t)/users\_close(t-1)$), *prc* (closing price), *size* (market cap in millions), *ret* (daily return), *openret* (overnight return), *dayret* (daytime return), *daily_buys* (number of TAQ daily retail buys), *daily_sells* (number of TAQ daily retail sells), *net_buys* ($daily\_buys - daily\_sells$), *taq_retimb* ($net\_buys/(daily\_buys + daily\_sells)$).

variable	N	mean	sd	min	p25	p50	p75	max
users_close	4,884	2,487.02	7,573.30	116.00	353.00	774.50	1,914.50	154,351.00
users_last	4,884	2,605.28	7,887.32	118.00	367.00	810.00	2,007.50	156,826.00
userchg	4,884	1,103.72	3,514.05	16.00	119.00	288.50	803.00	85,193.00
userratio	4,884	1.99	1.69	1.10	1.37	1.56	1.98	44.71
prc	4,712	163.89	6,973.02	0.12	3.34	8.95	21.39	341,000.00
size ($mil)	4,299	2,231.98	11,532.14	0.11	45.29	375.29	1,229.05	468,894.20
ret (%)	4,711	14.02	52.58	-91.79	-7.10	4.88	20.56	874.84
openret (%)	4,707	10.99	39.19	-88.60	-0.94	2.04	11.61	563.90
dayret (%)	4,710	3.43	30.68	-74.69	-7.79	-0.06	8.02	595.19
daily_buys	4,675	3,498.46	9,536.55	0.00	252.00	774.00	2,669.00	162,678.00
daily_sells	4,675	2,710.01	7,108.48	0.00	205.00	615.00	2,149.00	110,145.00
net_buys	4,675	788.44	2,962.92	-8,873.00	7.00	98.00	487.00	56,504.00
taq_retimb	4,675	0.11	0.17	-0.75	0.01	0.09	0.20	1.00

Table 3: Concentration of Retail Buying and Selling

The sample consists of stocks with a measure of Robinhood user changes and net retail buying in TAQ. Statistics are based on averages across days. In Panel A, HH_buy is the Herfindahl–Hirschman index for stocks with net buying (sum of squared share of buying) and $Top10\_buy$ is the percentage of all net buying in the 10 stocks with the highest level of net buying. In Panel B, HH_sell is the Herfindahl–Hirschman index for stocks with net selling (sum of squared share of selling) and $Top10\_sell$ is the percentage of all net selling in the 10 stocks with the highest level of net selling. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	Robinhood User Changes	TAQ Retail Trades	Difference (RH - TAQ)
Panel A: Mean Daily Concentration Measures among Stocks with Net Buying			
HH_buy (bps)	246.952***	110.820***	136.133***
	(13.091)	(5.575)	(11.433)
$Top10\_buy$ (%)	34.621***	23.560***	11.061***
	(0.412)	(0.315)	(0.361)
Panel B: Mean Daily Concentration Measures among Stocks with Net Selling			
HH_sell (bps)	172.760***	48.566***	124.194***
	(20.069)	(2.357)	(20.055)
$Top10\_sell$ (%)	25.370***	14.076***	11.294***
	(0.369)	(0.218)	(0.384)

Table 4: Determinants of Robinhood Herding Indicator Variable (Top 0.5% of Positive Percentage User Change)

The table examines the determinants that predict the Robinhood herding indicator variable (top 0.5% of positive user change ratio on day t and a minimum of 100 users on day $t-1$) using a linear probability model. $rh\_herd(t-1)$ is the lagged Robinhood herding indicator. Extreme absolute return $(t-1)$ is an indicator that equals one if the absolute return is ranked in the top 20 on day $t-1$. Abnormal Vol $(t-1)$ is the logarithm of the ratio of stock market volume on day $t-1$ to the average volume from day $t-21$ to $t-2$. User Change $(t-1)$ is the change in Robinhood users from day $t-2$ to day $t-1$. $ln(Users(t-1))$ is the logarithm of Robinhood users before market closes on day $t-1$. Abnormal SVI $(t-1)$ is the logarithm of the ratio of Google search volume on day $t-1$ to the average Google search volume from day $t-21$ to $t-2$. Abnormal News $(t-1)$ is the logarithm of the ratio of news article count on day $t-1$ to the average news article count from day $t-21$ to $t-2$. Earnings Announcement $(t-1)$ is an indicator that equals one if the firm has an earnings announcement on day $t-1$. Robust standard errors are clustered on day and stock level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)
Dep var:		$rh\_herd(t)$		
$rh\_herd(t-1)$	0.106***	0.103***	0.102***	0.102***
	(0.005)	(0.005)	(0.005)	(0.005)
Extreme absolute return $(t-1)$	0.0505***	0.0494***	0.0485***	0.0486***
=1: top 20 absret; =0: otherwise	(0.003)	(0.003)	(0.003)	(0.003)
Abnormal Vol $(t-1)$	0.000551***	0.000453***	0.000368***	0.000348***
$= ln(Vol(t-1)/AvgVol(t-21,t-2))$	(0.000)	(0.000)	(0.000)	(0.000)
User Change $(t-1)$ (in 000s)		0.00806***	0.00738***	0.00733***
		(0.001)	(0.001)	(0.001)
$ln(Users(t-1))$		0.000193***	0.000170***	0.000166***
		(0.000)	(0.000)	(0.000)
Abnormal SVI $(t-1)$			0.000149***	0.000142***
$= ln(SVI(t-1)/AvgSVI(t-21,t-2))$			(0.000)	(0.000)
Abnormal News $(t-1)$			0.00195***	0.00107***
$= ln(News(t-1)/AvgNews(t-21,t-2))$			(0.000)	(0.000)
Earnings Announcement $(t-1)$				0.00875***
				(0.001)
Observations	3792584	3792584	3792584	3792584
R-squared	0.022	0.022	0.023	0.023

Table 5: The Effect of Robinhood Outages on Percent Retail Trade

The dependent variable is the proportion of TAQ trades that are identified as retail trades per period. Outage is an indicator variable that takes a value of one at the time of an outage. In Panels B and C, Repair is an indicator variable that takes a value of one for the hour after systems are fully operational. In Panel C, Partial is an indicator variable that takes a value of one for the period when systems are partially restored. Column (1) presents results for all stocks; column (2) for the 50 most popular Robinhood stocks, and column (3) for the 50 high attention stocks (based on fitted values of model 3, Table 4). In Panel A, the Robinhood outage is for the full day on March 2, 2020, and observations are stock-hours. In Panel B, the time of the Robinhood outage is March 3, 2020, at 10:15 am with all systems back online sometime between 11 am and noon; observations are stock-hours. The dataset for Panels A and B spans February 18 to March 17. In Panel C, the time of the Robinhood outage is June 18, 2020, at 10:39 am, with systems improvement at 12:43 pm and fully restored at 1:08 pm. The dataset for Panel C spans June 4 to July 2. Outage is an indicator variable that takes a value of one for the time intervals between 11:35 am and 12:35 pm, June 18; observations are stock-five-minute periods. Robust standard errors are double clustered on day and ticker. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	All Stocks	50 Popular Stocks	50 High Attention Stocks
Panel A: March 2 Outage (all day)			
Outage	-0.723***	-5.227***	-4.948***
	(0.203)	(0.338)	(0.340)
Observations	1,090,382	8,395	8,319
R-squared	0.393	0.745	0.618
Day FE	NO	NO	NO
Ticker FE	YES	YES	YES
Time of Day FE	YES	YES	YES
Panel B: March 3 Outage (late morning)			
Outage	-1.720***	-6.610***	-5.122***
	(0.172)	(0.712)	(0.550)
Repair	1.581***	3.742***	1.652***
	(0.133)	(0.131)	(0.124)
Observations	1,038,326	7,997	7,821
R-squared	0.397	0.764	0.614
Day FE	YES	YES	YES
Ticker FE	YES	YES	YES
Time of Day FE	YES	YES	YES
Panel C: June 18 Outage (late morning)			
Outage	-0.678***	-3.042***	-0.752**
	(0.067)	(0.455)	(0.350)
Partial	0.476***	1.005***	0.973**
	(0.091)	(0.312)	(0.346)
Repair	0.731***	1.543***	0.458***
	(0.092)	(0.190)	(0.145)
Observations	9,443,439	81,814	64,652
R-squared	0.231	0.621	0.176
Day FE	YES	YES	YES
Ticker FE	YES	YES	YES
Time of Day FE	YES	YES	YES

Table 6: The Effect of Top Movers on Robinhood (RH) Intraday User Change vs. TAQ Intraday Net Retail Buying

The table examines how the top mover rankings affect Robinhood (RH) intraday user change and TAQ intraday net retail buying. The rankings are sorted for stocks with market cap above $300 million at the market open of day t. The stocks are sorted on absolute overnight return (columns (1) and (2)) and absolute daily return (columns (3) and (4)), respectively. The sample requires both Robinhood user change and TAQ net retail buying available and only includes the top 20 stocks for each day. RH Intraday User Change measures the change in RH users from first time stamp that excludes market open trades to the last time stamp before market closes. TAQ Intraday Net Retail Buying is TAQ retail buying minus TAQ retail selling (with short trades removed following Boehmer and Song (2020)) during the market trading hours. Top mover score assigns a score for each rank, with 20 for the highest absolute return and 1 for the 20th highest. Negative return is an indicator variable that equals one if the stock return is negative. Regressions include day fixed effects. Robust standard errors are clustered on day level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)
Dep var:	RH Intraday User Change (t)	TAQ Intraday Net Buy (t)	RH Intraday User Change (t)	TAQ Intraday Net Buy (t)
Top mover score is sorted on:	Absolute Overnight Return (t)		Absolute Return (t)	
Top mover score	7.198***	13.41***	11.95***	17.89***
=20: highest absret; =1: 20th highest absret	(0.510)	(1.701)	(0.539)	(1.671)
Negative return	46.49***	-133.2***	2.980	-86.90***
=1: top mover return is negative; =0, otherwise	(9.913)	(29.950)	(10.252)	(28.705)
Top mover score X Negative return	3.462***	-6.591***	-1.472**	-6.245***
	(0.736)	(2.369)	(0.724)	(2.222)
Day FE	Yes	Yes	Yes	Yes
Observations	10247	10247	10342	10342
R-squared	0.233	0.211	0.272	0.192

Table 7: Regression Discontinuity in Robinhood (RH) Intraday User Change for Top Mover Stocks

This table estimates a sharp RD regression that exploits the discontinuity in Robinhood (RH) intraday user changes around the market cap cutoff of $300 million. Specifically, we use Robinhood (RH) intraday user changes at day (t) as the dependent variable, and include an indicator that equals one for stocks that have market cap at the market open of day t greater than $300 million. We include different polynomial functions of market cap as controls. Our analysis varies sample bandwidth from Panel A to Panel D. Panel A uses a sample bandwidth of 50 million (i.e. market cap $\in$ [$250 million, $350 million]). For stocks with market cap above $300 million, we select stocks that rank top 20 by absolute day-t overnight returns among all stocks with market cap above $300 million. For stocks with market cap below $300 million, we include matched stocks with absolute day-t overnight returns close to stocks with market cap above $300 million. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)
Dep var:		Intraday RH User Change	
Panel A: Sample bandwidth = 50m			
Larger than 300m	95.71**	115.8**	247.3***
	(39.472)	(53.933)	(78.782)
Polynomial order, N	1	2	3
Observations	1332	1332	1332
R-squared	0.013	0.014	0.020
Panel B: Sample bandwidth = 75m			
Larger than 300m	90.57**	109.8*	97.18
	(37.123)	(61.131)	(80.490)
Polynomial order, N	1	2	3
Observations	2068	2068	2068
R-squared	0.006	0.006	0.008
Panel C: Sample bandwidth = 100m			
Larger than 300m	77.55**	131.0**	174.8**
	(33.300)	(53.648)	(72.724)
Polynomial order, N	1	2	3
Observations	2782	2782	2782
R-squared	0.011	0.011	0.012
Panel D: Sample bandwidth = 125m			
Larger than 300m	49.13*	102.3**	147.7**
	(26.718)	(41.947)	(59.496)
Polynomial order, N	1	2	3
Observations	3406	3406	3406
R-squared	0.016	0.016	0.016

Table 8: Event Time Abnormal Returns

The table reports the abnormal returns around Robinhood user herding events. Abnormal returns (*AR*) are computed as the raw return minus the CRSP value-weighted average return. Abnormal returns are averaged across all events. Buy-and-hold abnormal returns (*BHAR*) are computed as the product of one plus the stock's return through event day t less the product of one plus the market return for the same period. Standard errors are computed by clustering on event day. *%Positive* is the percent of returns that are positive. Herding events are defined as the top 0.5% of stocks with positive user change ratio on day 0 and a minimum of 100 users on the prior day. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Event Day	AR	Std. Err.	% Positive	BHAR	Std. Err.	% Positive
			Pre-Event			
-10	-0.08%	0.14%	47%	-0.08%	0.14%	47%
-9	-0.05%	0.15%	46%	-0.02%	0.27%	46%
-8	0.38%	0.30%	47%	0.22%	0.41%	46%
-7	-0.06%	0.21%	45%	-0.01%	0.50%	45%
-6	-0.28%**	0.13%	46%	-0.48%	0.42%	44%
-5	0.02%	0.15%	48%	-0.47%	0.45%	45%
-4	0.39%*	0.20%	48%	-0.10%	0.50%	46%
-3	0.51%**	0.21%	48%	0.52%	0.59%	47%
-2	0.40%*	0.22%	49%	0.81%	0.60%	48%
-1	4.59%***	0.52%	56%	5.60%***	0.87%	51%
0	13.95%***	0.92%	63%	22.16%***	1.73%	58%
			Post-Event			
1	-1.23%***	0.24%	42%	-1.23%***	0.24%	42%
2	-0.85%***	0.18%	42%	-2.11%***	0.31%	40%
3	-0.43%***	0.16%	44%	-2.74%***	0.29%	38%
4	-0.35%**	0.15%	43%	-3.15%***	0.31%	37%
5	-0.32%**	0.14%	44%	-3.55%***	0.32%	37%
6	-0.37%***	0.13%	44%	-3.99%***	0.34%	37%
7	-0.14%	0.14%	46%	-4.17%***	0.37%	36%
8	0.17%	0.17%	45%	-4.13%***	0.39%	36%
9	0.07%	0.13%	45%	-4.19%***	0.40%	36%
10	0.15%	0.16%	45%	-4.15%***	0.40%	37%
11	-0.03%	0.14%	42%	-4.33%***	0.40%	36%
12	-0.03%	0.12%	46%	-4.42%***	0.41%	36%
13	-0.06%	0.13%	44%	-4.51%***	0.43%	36%
14	-0.09%	0.14%	45%	-4.68%***	0.44%	35%
15	-0.15%	0.10%	45%	-4.87%***	0.45%	35%
16	0.03%	0.11%	46%	-4.83%***	0.47%	35%
17	-0.04%	0.14%	45%	-4.76%***	0.56%	35%
18	0.18%	0.22%	46%	-4.71%***	0.56%	35%
19	-0.01%	0.11%	46%	-4.80%***	0.57%	35%
20	0.15%	0.16%	46%	-4.74%***	0.58%	35%

Table 9: Calendar Portfolio Returns of Herding Events

The dependent variable is the dollar weighted average daily return over the risk-free rate (%). For each *rh_herd* event (top 0.5% of positive user change ratio on day t and a minimum of 100 users on day $t-1$), $1/Price$ shares are purchased at the end of the herding day. These stocks are held for five days before being liquidated. Dollar weighted average daily returns are a dollar weighted average of all stocks held based on the position value at the end of the prior day. Returns are over the entire period (columns (1)-(3)), prior to March 2020 (columns (4)-(6)), or March 2020 and after (columns (7)-(9)). The key estimation is the constant, $Alpha$. Control variables include excess market returns ($R_m - R_f$), small-minus-big factor (*SMB*), high-minus-low factor (HML), momentum factor (MOM), robust-minus-weak operating profitability factor (RMW), and conservative-minus-aggressive factor (*CMA*). Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep var:	Daily excess return on dollar-weighted calendar portfolio of herding events (%)								
	Entire Period			Prior to March 2020			March 2020 and after		
Alpha	-0.612***	-0.569***	-0.552***	-0.535***	-0.526***	-0.511***	-0.938***	-0.818***	-0.777***
	(0.112)	(0.110)	(0.108)	(0.118)	(0.118)	(0.118)	(0.305)	(0.284)	(0.273)
Mkt_Rf	0.811***	0.709***	0.660***	0.646***	0.512***	0.437***	0.883***	0.725***	0.699***
	(0.111)	(0.120)	(0.127)	(0.120)	(0.130)	(0.135)	(0.138)	(0.155)	(0.156)
SMB		0.689***	0.446**		0.467	0.332		0.719**	0.399
		(0.218)	(0.219)		(0.308)	(0.304)		(0.317)	(0.384)
HML		0.103	0.355		-0.421*	-0.148		0.298	0.533
		(0.176)	(0.217)		(0.231)	(0.269)		(0.311)	(0.394)
MOM		-0.118	-0.229		-0.511**	-0.599**		0.050	-0.085
		(0.158)	(0.166)		(0.233)	(0.238)		(0.232)	(0.266)
RMW			-0.968***			-0.981***			-0.961
			(0.301)			(0.337)			(0.595)
CMA			-0.773*			-0.703			-0.858
			(0.424)			(0.520)			(0.749)
Observations	555	555	555	448	448	448	107	107	107
R-squared	0.190	0.234	0.256	0.056	0.079	0.098	0.407	0.481	0.496

Table 10: Regression of Daily Returns on Lagged Robinhood Herding Indicator (Top 0.5% Percentage User Change)

The dependent variable is the daily stock return (%) winsorized at the 0.1% level (*ret*). In columns (1) to (3), $rh\_herd(t)$ is an indicator variable that equals one if the percentage change in users is in the top 0.5% for stocks with positive user changes on day t and a minimum of 100 users on day $t-1$. In columns (4) to (6) (or columns (7) to (9)), the $rh\_herd(t)$ indicator variable equals one if the prior conditions are met and the overnight return from the close on day $t-1$ to the open on day t is positive (negative). Control variables include retail order imbalance from TAQ (*taq_retimb*), lagged returns (*ret*), lags of an indicator variable if the *rh_herd* measure is missing (and *rh_herd* is set equal to zero), and day fixed effects. 5-day AR (%) is the sum of the coefficients on the five lags of *rh_herd*. Robust standard errors clustered by day are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep var:					$ret(t)$ (%)				
Herding Events:		All Events			Overnight Return > 0			Overnight Return < 0	
$rh\_herd(t-1)$	-1.336***	-1.039***	-1.110***	-1.756***	-1.134***	-1.276***	-0.514**	-0.970***	-0.894***
	(0.160)	(0.171)	(0.174)	(0.215)	(0.257)	(0.260)	(0.209)	(0.226)	(0.229)
$rh\_herd(t-2)$	-0.758***	-0.813***	-0.798***	-1.150***	-1.235***	-1.212***	-0.152	-0.140	-0.152
	(0.129)	(0.141)	(0.141)	(0.175)	(0.221)	(0.220)	(0.175)	(0.202)	(0.203)
$rh\_herd(t-3)$	-0.298**	-0.191	-0.183	-0.377**	-0.148	-0.127	-0.268	-0.431**	-0.449**
	(0.122)	(0.134)	(0.132)	(0.160)	(0.203)	(0.199)	(0.194)	(0.208)	(0.207)
$rh\_herd(t-4)$	-0.318***	-0.266**	-0.266**	-0.310**	-0.169	-0.178	-0.388**	-0.557***	-0.534***
	(0.115)	(0.128)	(0.128)	(0.144)	(0.183)	(0.182)	(0.163)	(0.181)	(0.181)
$rh\_herd(t-5)$	-0.231**	-0.286**	-0.273**	-0.303**	-0.344*	-0.326*	-0.167	-0.243	-0.237
	(0.117)	(0.131)	(0.131)	(0.151)	(0.190)	(0.190)	(0.166)	(0.171)	(0.171)
$ret(t-1)$		-0.046***	-0.037***		-0.045***	-0.036***		-0.047***	-0.038***
		(0.012)	(0.012)		(0.012)	(0.012)		(0.012)	(0.012)
$ret(t-2)$		-0.002	-0.002		-0.001	-0.001		-0.003	-0.003
		(0.012)	(0.012)		(0.012)	(0.012)		(0.012)	(0.012)
$ret(t-3)$		-0.021*	-0.021*		-0.021*	-0.021*		-0.022*	-0.021*
		(0.012)	(0.012)		(0.012)	(0.012)		(0.012)	(0.012)
$ret(t-4)$		-0.017*	-0.017*		-0.017*	-0.017*		-0.018*	-0.017*
		(0.010)	(0.010)		(0.010)	(0.010)		(0.010)	(0.010)
$ret(t-5)$		-0.005	-0.005		-0.005	-0.005		-0.005	-0.006
		(0.010)	(0.010)		(0.010)	(0.010)		(0.010)	(0.010)
$taq\_retimb(t-1)$			0.043***			0.043***			0.042***
			(0.008)			(0.008)			(0.008)
$taq\_retimb(t-2)$			0.012			0.012			0.010
			(0.008)			(0.008)			(0.008)
$taq\_retimb(t-3)$			0.000			0.000			-0.000
			(0.008)			(0.009)			(0.009)
$taq\_retimb(t-4)$			0.009			0.009			0.009
			(0.008)			(0.008)			(0.008)
$taq\_retimb(t-5)$			0.009			0.009			0.008
			(0.008)			(0.008)			(0.008)
Observations	3,656,926	3,652,401	3,312,553	3,656,926	3,652,401	3,312,553	3,656,926	3,652,401	3,312,553
R-squared	0.196	0.199	0.205	0.196	0.199	0.205	0.196	0.198	0.205
Days	550	550	550	550	550	550	550	550	550
5-day AR (%)	-2.942***	-2.595***	-2.630***	-3.896***	-3.029***	-3.120***	-1.490***	-2.341***	-2.265***
Std. Err.	(0.322)	(0.350)	(0.348)	(0.435)	(0.523)	(0.520)	(0.423)	(0.447)	(0.451)

Table 11: Regression of Daily Returns on Lagged Robinhood Herding Indicator and TAQ Retail Herding Indicator

The dependent variable is the daily stock return (%) winsorized at the 0.1% level (*ret*). $rh\_herd(t)$ is an indicator variable that equals one if the percentage change in users is in the top 0.5% for stocks with positive user changes on day t and a minimum of 100 users on day $t-1$. On each day, we identify the same number of herding events (N) using the TAQ retail trade data to construct $taq\_herd(t)$, an indicator variable that equals one if the stock is among the N stocks with the greatest abnormal retail volume within the top quintile of standardized retail order imbalance on day t. In columns (3) and (4), we include an indicator variable for whether TAQ retail order imbalance is positive (*taqpos*) and its interaction with *rh_herd*. Control variables include retail order imbalance from TAQ (*taq_retimb*), lagged returns (*ret*), lags of an indicator variable if the *rh_herd* measure is missing (and *rh_herd* is set equal to zero), and day fixed effects. 5-day AR (%) is the sum of the coefficients on the five lags of *rh_herd* (or *taq_herd*). Robust standard errors clustered by day are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:			$ret(t)$ (%)			
$rh\_herd(t-1)$	-1.336***		-0.549**	-0.514**	-1.001***	-0.838***
	(0.160)		(0.234)	(0.241)	(0.160)	(0.170)
$rh\_herd(t-2)$	-0.758***		-0.306	-0.358*	-0.550***	-0.587***
	(0.129)		(0.194)	(0.193)	(0.138)	(0.144)
$rh\_herd(t-3)$	-0.298**		-0.099	-0.086	-0.232*	-0.166
	(0.122)		(0.191)	(0.192)	(0.126)	(0.128)
$rh\_herd(t-4)$	-0.318***		-0.356**	-0.356**	-0.369***	-0.337***
	(0.115)		(0.159)	(0.162)	(0.112)	(0.120)
$rh\_herd(t-5)$	-0.231**		-0.196	-0.204	-0.246**	-0.277**
	(0.117)		(0.182)	(0.189)	(0.118)	(0.127)
$taq\_herd(t-1)$		-1.220***			-0.896***	-0.794***
		(0.154)			(0.151)	(0.182)
$taq\_herd(t-2)$		-0.772***			-0.485***	-0.536***
		(0.120)			(0.120)	(0.137)
$taq\_herd(t-3)$		-0.180			0.001	0.147
		(0.113)			(0.107)	(0.125)
$taq\_herd(t-4)$		0.050			0.173	0.252*
		(0.109)			(0.106)	(0.129)
$taq\_herd(t-5)$		-0.087			-0.006	-0.022
		(0.100)			(0.106)	(0.124)
$rh\_herd(t-1) \times taqpos(t-1)$			-1.021***	-0.759**		
			(0.300)	(0.310)		
$rh\_herd(t-2) \times taqpos(t-2)$			-0.589**	-0.567**		
			(0.249)	(0.259)		
$rh\_herd(t-3) \times taqpos(t-3)$			-0.260	-0.129		
			(0.239)	(0.263)		
$rh\_herd(t-4) \times taqpos(t-4)$			0.051	0.121		
			(0.213)	(0.223)		
$rh\_herd(t-5) \times taqpos(t-5)$			-0.042	-0.085		
			(0.212)	(0.218)		

$rh\_herd(t-1) \times taq\_herd(t-1)$					-0.165	-0.049
					(0.399)	(0.407)
$rh\_herd(t-2) \times taq\_herd(t-2)$					-0.252	-0.230
					(0.329)	(0.334)
$rh\_herd(t-3) \times taq\_herd(t-3)$					-0.262	-0.239
					(0.316)	(0.321)
$rh\_herd(t-4) \times taq\_herd(t-4)$					0.015	0.020
					(0.267)	(0.273)
$rh\_herd(t-5) \times taq\_herd(t-5)$					0.066	0.044
					(0.262)	(0.265)
Observations	3,656,926	3,656,926	3,656,926	3,312,553	3,656,926	3,312,553
R-squared	0.196	0.196	0.196	0.205	0.196	0.205
Lagged Control Variables:						
ret	NO	NO	NO	YES	NO	YES
taq_retimb	NO	NO	NO	YES	NO	YES
taqpos	NO	NO	YES	YES	NO	NO
RH 5-day AR (%)	-2.942***		-1.506***	-1.518***	-2.397***	-2.204***
RH Std. Err.	(0.322)		(0.456)	(0.461)	(0.310)	(0.336)
TAQ 5-day AR (%)		-2.208***			-1.213***	-0.954***
TAQ Std. Err.		(0.347)			(0.321)	(0.326)

Table 12: Subsample of Post-Herding Return Patterns

The table presents the 5-day abnormal return (%) from specifications (1) to (3) of Table 10 for various subsamples. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)
Panel A: Full Sample			
5-day AR (%)	-2.942***	-2.595***	-2.630***
Std. Err.	(0.322)	(0.350)	(0.348)
Panel B: Quote Midpoint Returns			
5-day AR	-2.797***	-2.378***	-2.408***
Std. Err.	(0.322)	(0.357)	(0.354)
Panel C: Only stock with prices > \$5			
5-day AR	-0.851***	-0.756***	-0.749**
Std. Err.	(0.242)	(0.266)	(0.265)
Panel D: Small Cap (< \$1 billion in market cap)			
5-day AR	-4.250***	-3.766***	-3.835***
Std. Err.	(0.418)	(0.434)	(0.432)
Panel E: Large Cap (> \$1 billion in market cap)			
5-day AR	0.14	-0.0348	-0.0508
Std. Err.	(0.534)	(0.517)	(0.521)
Panel F: Other than Common Stocks			
5-day AR	-3.174***	-2.895***	-2.940***
Std. Err.	(0.583)	(0.669)	(0.667)
Panel G: Post-Covid (after March 13, 2020)			
5-day AR	-4.581***	-2.876***	-3.073***
Std. Err.	(0.680)	(0.750)	(0.743)
Panel H: Pre-Covid (before March 13, 2020)			
5-day AR	-2.221***	-2.239***	-2.231***
Std. Err.	(0.337)	(0.327)	(0.330)

Table 13: Regression of Abnormal Short Trading on Returns and Top Robinhood Changes

The dependent variable is abnormal short volume. Short volume is measured using data from FINRA which provides daily short trade data as noted by Boehmer and Song (2020). We compute abnormal short volume on day t as the ratio of short trades on day t to the prior 20-day ($[t-20, t-1]$) average of the number of short trades. The top 0.1% of observations are winsorized. The key independent variables are *rh_chgratio*, which is the percentage change in users from day $t-1$ to t, and *rh_herd*, which is an indicator variable that equals one if the percentage change in users is in the top 0.5% for stocks with positive user changes on day t and a minimum of 100 users on day $t-1$. Control variables include the excess return of day t, dummy variables representing the return ranks (*ret_rank*) of day t, and abnormal news coverage which is the logarithm of the ratio of news article count on day t to the average news article count from day $t-20$ to $t-1$. Coefficients are in percent. Standard errors are computed using Fama-Macbeth (1973). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)
Dep var:	Abnormal Short Volume (%)			
rh_chgratio	637.14***		551.10***	
	(12.75)		(11.10)	
rh_herd		854.65***		686.57***
		(14.5)		(17.25)
Excess return	499.54***	498.83***		
	(16.32)	(15.95)		
Abnormal News	6.64***	7.17***	6.37***	6.85***
	(0.13)	(0.14)	(0.12)	(0.13)
*Ret_rank*_1			858.14***	1049.44***
			(47.95)	(36.36)
*Ret_rank*_2			704.80***	818.70***
			(44.31)	(34.22)
*Ret_rank*_3			568.94***	670.82***
			(68.35)	(32.64)
*Ret_rank*_4			583.67***	635.83***
			(35.00)	(31.91)
*Ret_rank*_5			534.46***	554.07***
			(30.95)	(30.50)
*Ret_rank*_6			465.52***	480.40***
			(26.90)	(26.95)
*Ret_rank*_7			410.78***	482.44***
			(39.91)	(28.08)
*Ret_rank*_8			371.97***	398.30***
			(27.60)	(25.01)
*Ret_rank*_9			381.49***	395.11***
			(26.37)	(26.27)
*Ret_rank*_10			361.47***	385.87***
			(24.36)	(24.53)
*Ret_rank*_$11-25$			269.46***	279.53***
			(8.20)	(8.55)
*Ret_rank*_$26-50$			145.48***	152.75***
			(4.30)	(4.56)
*Ret_rank*_$51-100$			83.45***	86.87***
			(2.31)	(2.45)
Observations	2,681,173	2,681,173	2,681,173	2,681,173
Avg. R-squared	0.162	0.146	0.271	0.257

Internet Appendix

"Attention-Induced Trading and Returns: Evidence from Robinhood Users"

A.1 Aggregate Investor Experience: Number of Investors Who Experience Profits or Losses

For another way to examine how investors perform during herding episodes, we look at the number of users who experience gains or losses. Even if aggregate returns are negative, which we find, it is possible that more investors experience gains rather than losses during the herding episode because losses might be heavily skewed.

To assess whether this is the case, for each event *j* we count the number of investors who record gains or losses during the 31-day event window, $\tau = -10, +20$. To do so, we compare the event day τ purchase price ($P_{i\tau}$) of new users ($\Delta u_{i\tau}$) to the price at the end of the event period ($P_{i,+20}$). The return is compared to the counterfactual of investing in an S&P 500 ETF ($P_{SP,\tau}; P_{SP,+20}$). We then count the number of users who have positive or negative profits relative to the counterfactual of investing in the S&P 500 ETF over the same period ($N^{buy}\_Pos_j$; $N^{buy}\_Neg_j$).

$$N^{buy}\_Pos_j = \sum_{\tau=-10}^{+20} \mathrm{UI}_{\mathrm{i\tau}} \Delta u_{i\tau} * I\left[\left(\frac{P_{i,+20}}{P_{i\tau}} - \frac{P_{SP,+20}}{P_{SP,\tau}}\right) > 0\right]. \quad (A.1)$$

$$N^{buy}\_Neg_j = \sum_{\tau=-10}^{+20} \mathrm{UI}_{\mathrm{i\tau}} \Delta u_{i\tau} * I\left[\left(\frac{P_{i,+20}}{P_{i\tau}} - \frac{P_{SP,+20}}{P_{SP,\tau}}\right) \leq 0\right]. \quad (A.2)$$

The indicator $UI_{i\tau}$ equals one on days when the change in users is positive, $\Delta u_{i\tau} > 0$, and $I[\cdot]$ is an indicator function that takes a value of one when the condition within the brackets is true.

These counts can be summed across events to calculate the percentage of users who experience profits across all events. We also calculate the percent of users who experience profits for each event:

$$PercPos_j^{buy} = \frac{N^{buy}\_Pos_j}{(N^{buy}\_Pos_j + N^{buy}\_Neg_j)}. \quad (A.3)$$

The calculations above ignore periods when we experience a decrease in users. For these periods, we conduct a similar analysis in which we compare event day τ sales price of users who sell to the price at the end of the event period and calculate $N^{sell}\_Pos_j$, $N^{sell}\_Neg_j$, and $PercPos_j^{sell}$:

$$N^{sell}\_Pos_j = \sum_{\tau=-10}^{+20} (1 - \mathrm{UI}_{\mathrm{i\tau}})|\Delta u_{i\tau}| * I\left[\left(\frac{P_{i,+20}}{P_{i\tau}} - \frac{P_{SP,+20}}{P_{SP,\tau}}\right) < 0\right]. \quad (A.4)$$

$$N^{sell}\_Neg_j = \sum_{\tau=-10}^{+20} (1 - \mathrm{UI}_{\mathrm{i\tau}})|\Delta u_{i\tau}| * I\left[\left(\frac{P_{i,+20}}{P_{i\tau}} - \frac{P_{SP,+20}}{P_{SP,\tau}}\right) \geq 0\right]. \quad (A.5)$$

$$PercPos_j^{sell} = \frac{N^{sell}\_Pos_j}{(N^{sell}\_Pos_j + N^{sell}\_Neg_j)}. \tag{A.6}$$

If $PercPos_j^{sell} = 0$, all users who sold during the event period would have been better off waiting to sell at the end of the event period (assuming a counterfactual investment in the S&P 500 ETF). We average these ratios across all stocks and report results from these analyses in Internet Appendix Figure A6.

For buys, we find results consistent with the previously reported investor experiences. Only 42% of Robinhood users experience positive returns from buying herding stocks. For sells, investors do better – 60% sell at a higher price than at the end of the period. While these results may suggest that Robinhood users break even, the number of buy events (approximately 18.7 million) far exceeds sell events (approximately 3.5 million).[1]

[1] The number of sell events is low because we exclude sales that happen at the end of the period. Most Robinhood users buy and never sell. Thus, this accounts for most of the users.

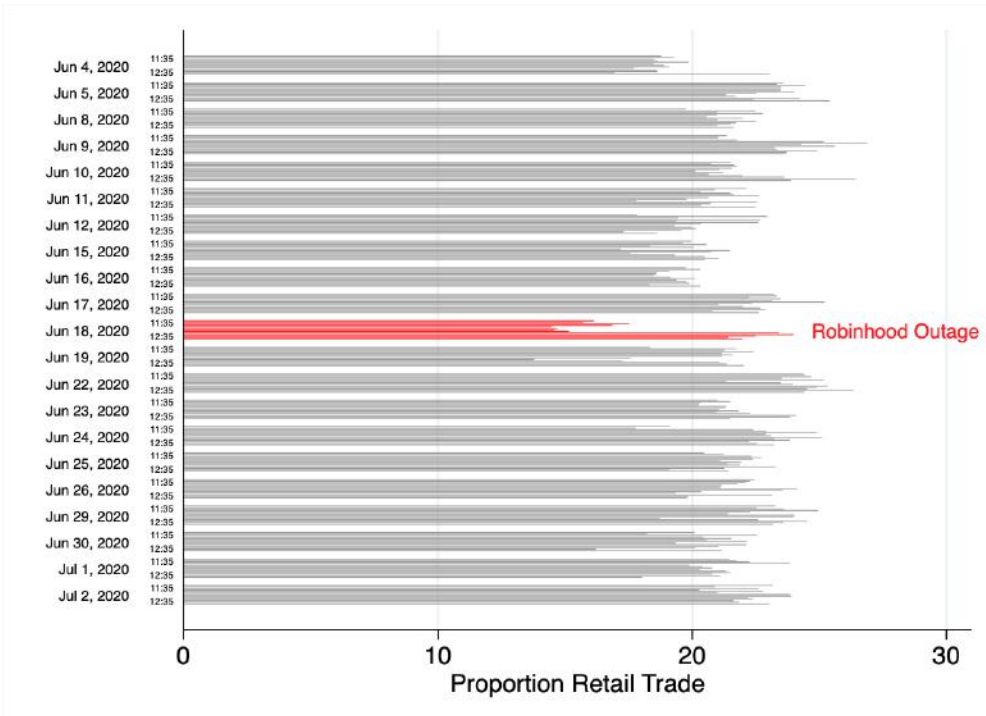


Figure A1: Percent Retail Trading in Five-Minute Intervals during June Outage

This figure depicts retail volume at five-minute intervals for the same period depicted in Figure 3, Panel D (11:35 am to 12:40 pm ET on June 18, 2020).

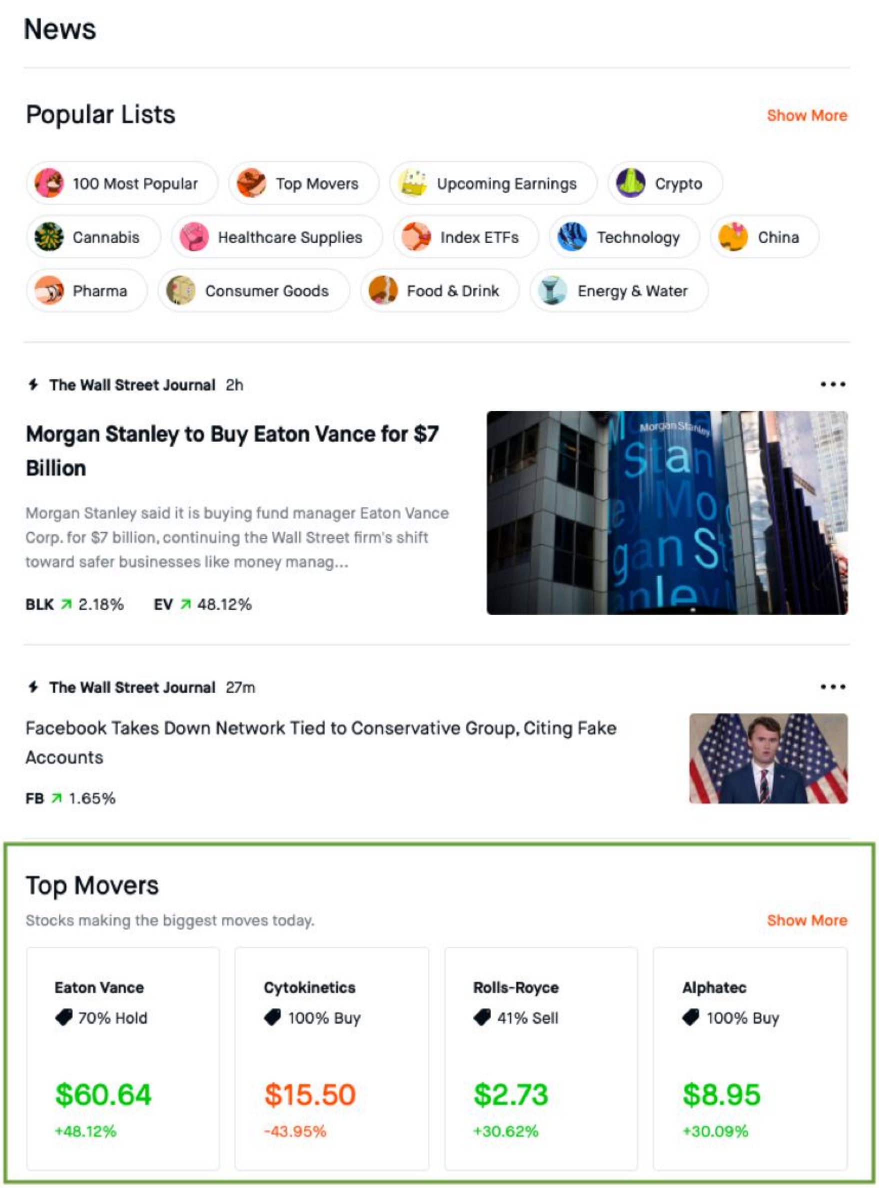


Figure A2: Screenshot of Robinhood User Interface

The figure presents a screenshot of the interface of the "News" section from Robinhood website on October 8, 2020. The “Top Movers” list is shown under the recent news. The initial screen presents the top four stocks with highest absolute return from the market close of the previous day (e.g., October 7, 2020 in this case). The full “Top Movers” list, which includes 20 stocks, could be accessed by clicking “Show More” or the “Top Movers” tab under “Popular Lists” on the top of the screenshot.

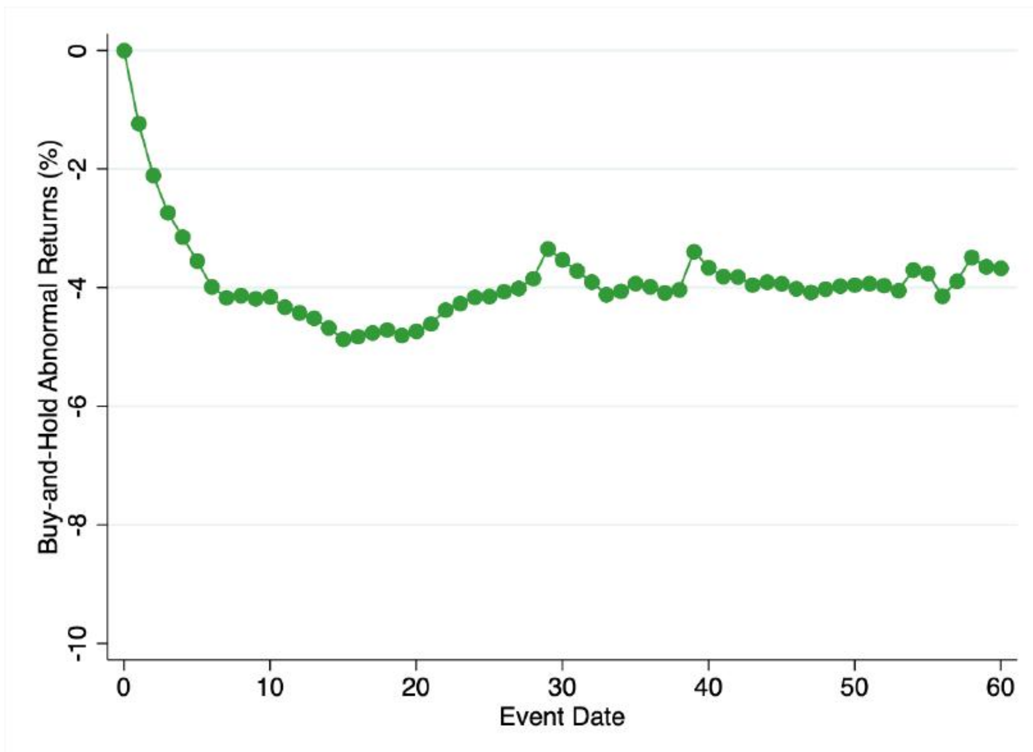
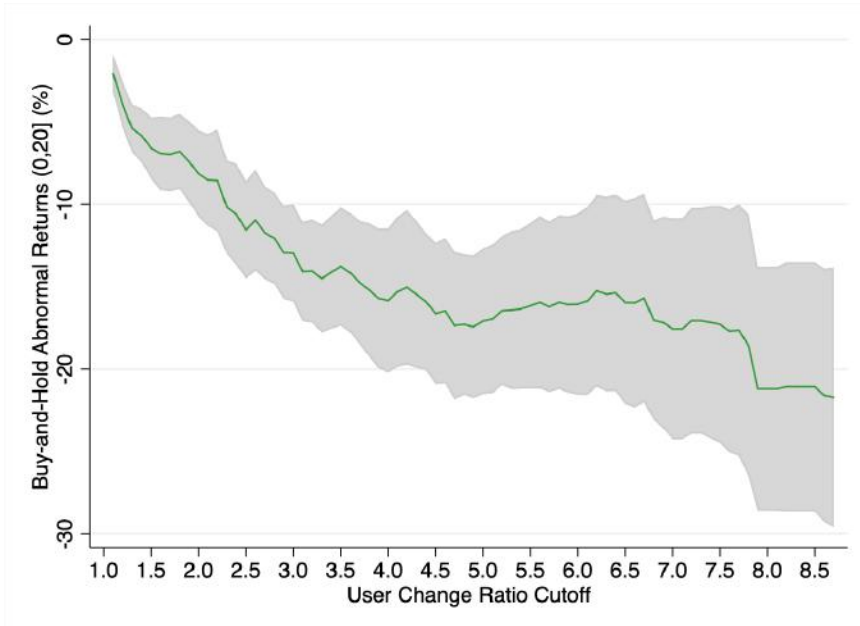


Figure A3: Returns Subsequent to Robinhood Herding Events: 60-day Horizon

The figure displays post-event mean BHARs starting from day 0 to day 60. Herding events are defined as the top 0.5% of stocks with positive user change ratio on day 0 and a minimum of 100 users on the prior day.

Panel A: Buy-and-hold Abnormal Returns (BHAR) and Daily User Change Ratio Cutoff



Panel B: Average Market Cap and Daily User Change Ratio Cutoff

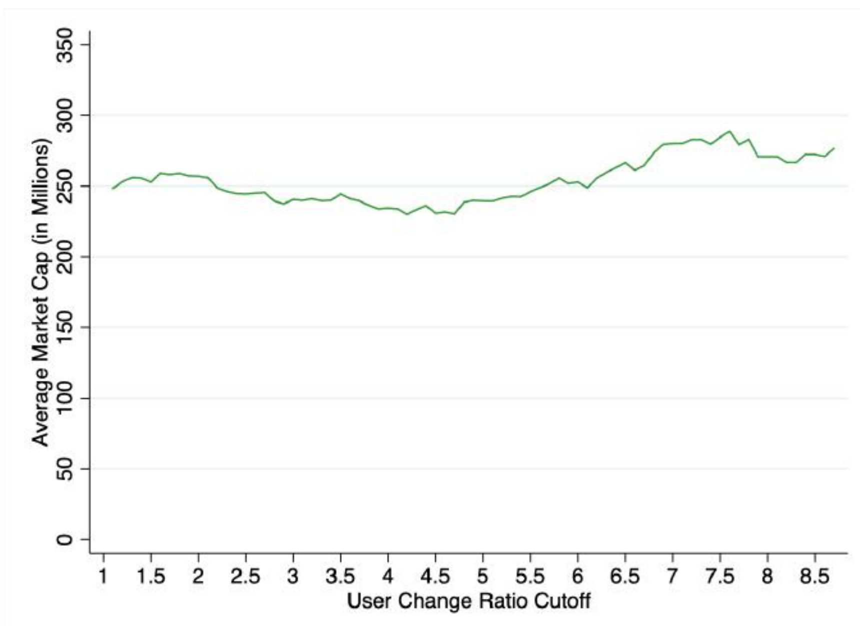


Figure A4: Price Reversal and Firm Size Characteristics (Market Cap $<$ \$1 Billion)

The figures show how 20-day buy-and-hold abnormal returns and average firm market cap vary with the daily user change ratio cutoff to identify the episodes for stocks with at least 100 Robinhood users. The analysis only includes firms with market cap less than \$1 Billion. The figure plots the 20-day buy-and-hold abnormal returns (%) against the user change ratio cutoff in Panel A and the average firm size (in Millions) in Panel B. The daily user change ratio cutoff varies from 1.1 to 8.5.

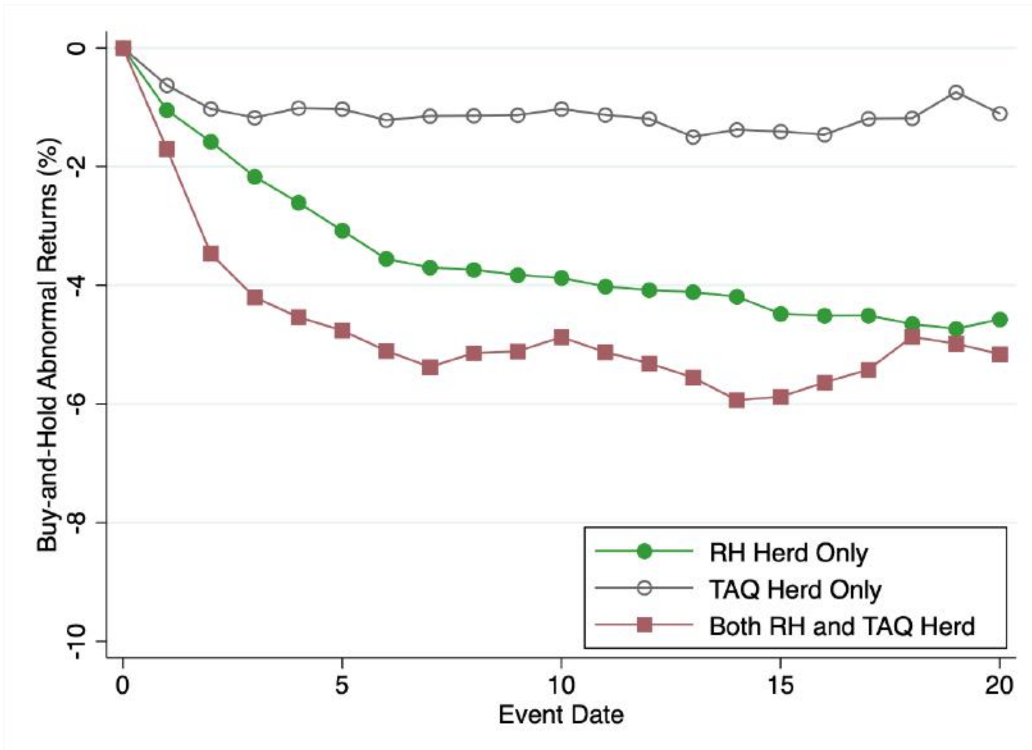


Figure A5: Returns Subsequent to Robinhood Herding Events vs. TAQ Herding Events

The figure displays post-event mean BHARs starting from event date 0. Robinhood herding events on day t are defined as the top 0.5% percentage change in users for stocks with positive user changes on day t and a minimum of 100 users on day $t-1$. On each day t, we identify the same number of TAQ herding events with the greatest abnormal retail volume within the top quintile of standardized retail order imbalance on day t. "RH Herding Only" includes all events that are classified as Robinhood herding events but not as TAQ herding events. "TAQ Herding Only" includes all events that are classified as TAQ herding events but not Robinhood herding events. "Both RH and TAQ Herd" includes all events that are classified as both.

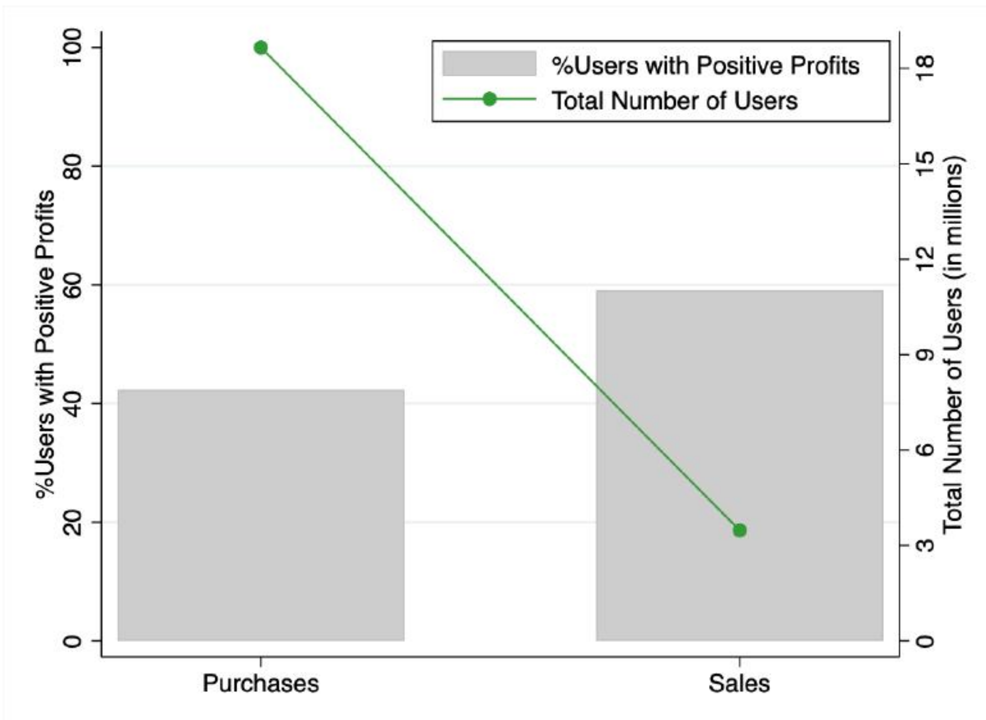


Figure A6: Aggregate Investor Experience: Percentage of Investors with Positive Profits

The figure presents the average percentage of investors who experience positive profits across herding episodes (left axis) and the total number of users who trade during the herding episodes (right axis). For each herding episode, we count the number of investors who record positive profits or negative profits during the 31-day event window [-10,20]. We separately consider purchases and sales. For purchases, we compute the profit as the ratio of the price at the end of the event period (day 20) to the purchase price relative to the ratio of corresponding prices of the S&P 500 ETF. For sales, we compute the profit as the ratio of the sales price to the price at the end of the event period (day 20) relative to the ratio of corresponding prices of the S&P 500 ETF.

Table A1: Overall Performance of Robinhood User Positions

The dependent variable is the daily excess return (portfolio return less risk-free rate, %) on a portfolio that mimics the Robinhood user experience. We assume that each new user buys $1/Price$ shares of stock; we update portfolio weights daily; and we assume decreases in user changes result in a sale of shares that is proportional to the percentage decrease in users. Returns are over the entire period (columns (1)-(3)), prior to March 2020 (columns (4)-(6)), or March 2020 and after (columns (7)-(9)). The key estimation is the constant, $Alpha$. Control variables include excess market returns ($R_m - R_f$), small-minus-big factor (SMB), high-minus-low factor (HML), momentum factor (MOM), robust-minus-weak operating profitability factor (RMW), and conservative-minus-aggressive factor (CMA). Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep var:	Daily Excess Return on Robinhood User Portfolio (%)								
	Entire Period			Prior to March 2020			March 2020 and after		
$Alpha$	0.006	0.012	0.021	-0.013	-0.016	-0.012	0.086	0.069	0.092
	(0.036)	(0.029)	(0.027)	(0.027)	(0.024)	(0.024)	(0.150)	(0.094)	(0.089)
$R_m - R_f$	1.084***	1.030***	0.986***	1.181***	1.084***	1.050***	1.044***	0.974***	0.955***
	(0.041)	(0.028)	(0.031)	(0.028)	(0.024)	(0.023)	(0.055)	(0.042)	(0.048)
SMB		0.467***	0.341***		0.281***	0.243***		0.555***	0.377***
		(0.062)	(0.059)		(0.081)	(0.081)		(0.079)	(0.092)
HML		-0.281***	-0.117*		-0.420***	-0.317***		-0.230	-0.092
		(0.061)	(0.067)		(0.058)	(0.065)		(0.115)	(0.130)
MOM		-0.380***	-0.405***		-0.222***	-0.243***		-0.449***	-0.512***
		(0.067)	(0.074)		(0.052)	(0.053)		(0.103)	(0.123)
RMW			-0.253**			-0.222***			-0.438*
			(0.106)			(0.072)			(0.232)
CMA			-0.605***			-0.300***			-0.545**
			(0.117)			(0.087)			(0.251)
Observations	552	552	552	445	445	445	107	107	107
R-squared	0.801	0.876	0.888	0.797	0.837	0.844	0.805	0.921	0.929

Table A2: Attention-Induced Return Effects sorted by Magnitude

Paper	Event	Horizon	Days	Sample Period	Trading Days	Abnormal Return	Events per Day
Robinhood Herding Events (this paper)	Robinhood user change ratio of 8.5 and on a base of 100 Robinhood users	day 1 to 20	20	May 2018 to Aug 2020	546	-19.64%	0.08
Robinhood Herding Events (this paper)	Robinhood user change ratio of 1.5 and at least 1000 new users	day 1 to 20	20	May 2018 to Aug 2020	546	-8.45%	1.65
Engelberg, Sasseville & Williams (2012)	Buy recommendations on Mad Money (the quintile of highest overnight returns)	day 1 to 50	50	Jul 2005 to Feb 2009	938-986	-6.94%	0.18
Robinhood Herding Events (this paper)	Top 0.5% of Daily Robinhood User Changes	day 1 to 20	20	May 2018 to Aug 2020	546	-4.74%	8.95
Barber & Loeffler (1993)	WSJ Pros in Monthly Dartboard Column with high Abnormal Volume	day 2 to 25	24	Oct 1988 to Oct 1990	525	-4.61%	0.07
Barber & Loeffler (1993)	WSJ Pros in Monthly Dartboard Column with high Abnormal Volume	day 2 to 5	4	Oct 1988 to Oct 1990	525	-2.23%	0.07
Cookson, Engelberg & Mullins (2020)	Declared Bull	day 1 to 10	10	2013 to 2019	1512	-1.85%	n.a.[(1)]
Cookson, Engelberg & Mullins (2020)	Declared Bull	day 1 to 5	5	2013 to 2019	1512	-1.41%	n.a.[(1)]
Tetlock (2011)	Stale momentum portfolio	day 1 to 10	10	Nov 1996 to Oct 2008	2980	-0.30%	66.00
Engelberg, Sasseville & Williams (2012)	Buy recommendations on Mad Money hosted by Jim Cramer	day 1 to 50	50	Jul 2005 to Feb 2009	938-986	insig	0.88
Da, Engelberg & Gao (2011)	Google Search Volume Index	week 5 to 52	232	Jan 2004 to Jun 2008	1134	insig	-

(1) The observations are StockTwits, which are not at the stock level and often duplicated for the event frequency.

Table A3: Regression Discontinuity Analysis: Return Matching Validity Test

This table estimates a sharp RD regression that exploits the discontinuity in the matching variable (absolute day-t overnight returns) around the market cap cutoff of $300 million. Specifically, we use the absolute overnight returns of day t as the dependent variable, and include an indicator that equals one for stocks that have market cap at the market open of day t greater than $300 million. We include different polynomial functions of market cap as controls. Our analysis uses a sample bandwidth of $50 million (i.e. market cap $\in$ [$250 million, $350 million]. For stocks with market cap above $300 million, we select stocks that rank top 20 by absolute day-t overnight returns among all stocks with market cap above $300 million. For stocks with market cap below $300 million, we include matched stocks with absolute day-t overnight returns close to stocks with market cap above $300 million. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)
Dep var:		Abs(Overnight return)	
Larger than 300m	0.00325	-0.00688	0.00235
	(0.011)	(0.019)	(0.021)
Sample bandwidth	50m	50m	50m
Polynomial order, N	1	2	3
N	1332	1332	1332
R-sq	0.014	0.016	0.016

Table A4: Placebo Test: Regression Discontinuity in Robinhood (RH) Intraday User Change for Top Mover Stocks

This table estimates a sharp RD regression that exploits the discontinuity in Robinhood (RH) intraday user changes around the market cap cutoff of $250 million. Specifically, we use Robinhood (RH) intraday user changes at day t as the dependent variable, and include an indicator that equals one for stocks that have market cap at the market open of day t greater than $250 million. We include different polynomial functions of market cap as controls. Our analysis uses a sample bandwidth of $50 million (i.e. market cap $\in$ [$200 million, $300 million]. For stocks with market cap above $250 million, we select stocks that rank top 20 by absolute day-t overnight returns among all stocks with market cap above $250 million. For stocks with market cap below $250 million, we include matched stocks with absolute day-t overnight returns close to stocks with market cap above $250 million. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)
Dep var:		Intraday RH User Change	
Larger than 250m	58.20	-11.11	-110.7
	(43.412)	(97.492)	(133.615)
Sample bandwidth	50m	50m	50m
Polynomial order, N	1	2	3
Observations	1666	1666	1666
R-squared	0.009	0.010	0.013

Table A5: Return Reversal Robustness: Regressions of Daily Returns on Lagged Robinhood Herding Indicator

Columns (1) and (2) present the results of the regression table from Table 10. Column (3) adds two indicator variables to column (1): top 20 most positive return stocks for the day ($top20pos$) and top 20 most negative return stocks for the day ($top20neg$). Column (4) adds two indicator variables to column (1): top 100 most positive return stocks ($top100pos$) and top 100 most negative return stocks ($top100neg$). Column (5) adds all four indicator variables to the model of column (1). Column (6) includes all indicator variables and lagged returns. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:			$ret(t)$ (%)			
$rh\_herd(t-1)$	-1.336***	-1.039***	-1.010***	-1.224***	-0.993***	-0.914***
	(0.160)	(0.171)	(0.152)	(0.155)	(0.152)	(0.155)
$rh\_herd(t-2)$	-0.758***	-0.813***	-0.558***	-0.742***	-0.568***	-0.586***
	(0.129)	(0.141)	(0.127)	(0.127)	(0.128)	(0.133)
$rh\_herd(t-3)$	-0.298**	-0.191	-0.256**	-0.276**	-0.253**	-0.207
	(0.122)	(0.134)	(0.122)	(0.120)	(0.122)	(0.127)
$rh\_herd(t-4)$	-0.318***	-0.266**	-0.300***	-0.310***	-0.301***	-0.257**
	(0.115)	(0.128)	(0.112)	(0.112)	(0.113)	(0.118)
$rh\_herd(t-5)$	-0.231**	-0.286**	-0.186	-0.241**	-0.204*	-0.219*
	(0.117)	(0.131)	(0.115)	(0.116)	(0.116)	(0.122)
$ret(t-1)$		-0.046***				-0.042**
		(0.012)				(0.018)
$ret(t-2)$		-0.002				0.004
		(0.012)				(0.017)
$ret(t-3)$		-0.021*				-0.027
		(0.012)				(0.018)
$ret(t-4)$		-0.017*				-0.027*
		(0.010)				(0.015)
$ret(t-5)$		-0.005				-0.005
		(0.010)				(0.015)
$top20pos(t-1)$			-1.293***		-1.005***	-0.562***
			(0.138)		(0.114)	(0.215)
$top20pos(t-2)$			-0.735***		-0.771***	-0.812***
			(0.133)		(0.109)	(0.195)
$top20pos(t-3)$			-0.223**		-0.141	0.140
			(0.114)		(0.100)	(0.205)
$top20pos(t-4)$			0.008		0.061	0.331*
			(0.108)		(0.087)	(0.174)
$top20pos(t-5)$			-0.117		-0.138	-0.093
			(0.102)		(0.084)	(0.179)

$top20neg(t-1)$			0.876***		0.328***	-0.004
			(0.124)		(0.100)	(0.158)
$top20neg(t-2)$			-0.208**		-0.186**	-0.151
			(0.096)		(0.083)	(0.157)
$top20neg(t-3)$			0.127		-0.014	-0.212
			(0.094)		(0.082)	(0.157)
$top20neg(t-4)$			-0.033		-0.091	-0.303**
			(0.086)		(0.076)	(0.144)
$top20neg(t-5)$			0.041		0.001	-0.051
			(0.086)		(0.076)	(0.138)
$top100pos(t-1)$				-0.476***	-0.276***	0.140
				(0.080)	(0.077)	(0.141)
$top100pos(t-2)$				-0.169**	-0.020	-0.075
				(0.080)	(0.078)	(0.127)
$top100pos(t-3)$				-0.144*	-0.107	0.164
				(0.076)	(0.077)	(0.125)
$top100pos(t-4)$				-0.080	-0.086	0.170*
				(0.068)	(0.067)	(0.103)
$top100pos(t-5)$				-0.027	0.002	0.037
				(0.072)	(0.072)	(0.102)
$top100neg(t-1)$				0.617***	0.569***	0.211**
				(0.074)	(0.072)	(0.102)
$top100neg(t-2)$				-0.065	-0.033	0.018
				(0.059)	(0.059)	(0.112)
$top100neg(t-3)$				0.111*	0.119**	-0.113
				(0.060)	(0.061)	(0.118)
$top100neg(t-4)$				0.021	0.047	-0.171*
				(0.056)	(0.057)	(0.098)
$top100neg(t-5)$				0.025	0.028	-0.017
				(0.054)	(0.055)	(0.108)
Observations	3,656,926	3,652,401	3,656,926	3,656,926	3,656,926	3,652,401
R-squared	0.196	0.199	0.197	0.197	0.197	0.199
5-day AR (%)	-2.942***	-2.595***	-2.311***	-2.793***	-2.318***	-2.183***
Std. Err.	(0.322)	(0.350)	(0.290)	(0.300)	(0.294)	(0.316)

Table A6: Event Time Abnormal Returns: TAQ Herding Events

The table reports the abnormal returns around TAQ herding events. On each day, we identify the same number of TAQ herding events as Robinhood herding events which have the greatest abnormal retail volume within the top quintile of standardized retail order imbalance. Abnormal returns (AR) are computed as the raw return minus the CRSP value-weighted average return. Abnormal returns are averaged across all events. Buy-and-hold abnormal returns (BHAR) are computed as the product of one plus the stock's return through event day t less the product of one plus the market return for the same period. Standard errors are computed by clustering on event day. % Positive is the percent of returns that are positive. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Event Day	AR	Std. Err.	% Positive	BHAR	Std. Err.	% Positive
			Pre-Event			
-10	-0.11%	0.11%	44%	-0.11%	0.11%	44%
-9	-0.18%*	0.11%	45%	-0.30%**	0.15%	44%
-8	-0.09%	0.11%	44%	-0.43%**	0.19%	43%
-7	-0.32%***	0.10%	44%	-0.79%***	0.23%	42%
-6	-0.15%	0.11%	45%	-0.96%***	0.26%	41%
-5	0.20%	0.17%	46%	-0.72%**	0.34%	42%
-4	-0.05%	0.11%	46%	-0.83%**	0.34%	42%
-3	0.30%**	0.14%	45%	-0.47%	0.38%	42%
-2	0.53%**	0.20%	47%	0.40%	0.60%	43%
-1	3.21%***	0.39%	52%	3.55%***	0.62%	46%
0	13.18%***	0.71%	66%	18.30%***	1.26%	58%
			Post-Event			
1	-0.92%***	0.24%	41%	-0.92%***	0.24%	41%
2	-0.71%***	0.18%	41%	-1.69%***	0.30%	39%
3	-0.18%	0.15%	43%	-2.00%***	0.31%	38%
4	0.13%	0.15%	44%	-1.97%***	0.31%	38%
5	0.01%	0.14%	44%	-2.04%***	0.33%	37%
6	-0.24%*	0.13%	44%	-2.27%***	0.34%	37%
7	-0.09%	0.13%	44%	-2.29%***	0.40%	37%
8	0.32%**	0.14%	45%	-2.23%***	0.37%	37%
9	0.16%	0.14%	43%	-2.21%***	0.38%	37%
10	0.17%	0.12%	45%	-2.07%***	0.41%	37%
11	-0.04%	0.11%	44%	-2.21%***	0.39%	36%
12	0.04%	0.12%	44%	-2.31%***	0.39%	36%
13	-0.21%*	0.11%	45%	-2.59%***	0.40%	36%
14	0.01%	0.10%	45%	-2.61%***	0.42%	36%
15	0.00%	0.11%	44%	-2.62%***	0.42%	35%
16	0.10%	0.10%	45%	-2.59%***	0.42%	36%
17	0.20%	0.13%	45%	-2.33%***	0.47%	36%
18	0.22%	0.21%	45%	-2.18%***	0.52%	36%
19	0.24%	0.15%	45%	-1.89%***	0.60%	35%
20	-0.13%	0.10%	44%	-2.20%***	0.57%	36%

Table A7: Calendar Portfolio Returns of Robinhood Herding Events vs. TAQ Herding Events

The dependent variable is the dollar weighted average daily returns of calendar portfolio of herding events over the risk-free rate (%). For each herding event observation, 1/Price shares are purchased at the end of the herding day. These stocks are held for five days before being liquidated. Dollar weighted average daily returns are a dollar weighted average of all stocks held based on the position value at the end of the prior day. Herding events are defined along two dimensions—Robinhood and TAQ herding events. Robinhood herding events are the top 0.5% of stocks with positive user changes on day t and a minimum of 100 users on day $t-1$. On each day, we identify the same number of TAQ herding events with the greatest abnormal retail volume within the top quintile of standardized retail order imbalance on day t. Column (1) form the portfolio based on all Robinhood herding events. Column (2) form the portfolio based on all TAQ herding events. Column (3) form the portfolio based on Robinhood herding events that are not TAQ herding events. Column (4) form the portfolio based on TAQ herding events that are not Robinhood herding events. Column (5) form the portfolio based on herding events that are classified as both. The key estimation is the constant, $Alpha$. Control variables include excess market returns (mkt_rf), small-minus-big factor (SMB), high-minus-low factor (HML), momentum factor (MOM), robust-minus-weak operating profitability factor (RMW), and conservative-minus-aggressive factor (CMA). Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)	(3)	(4)	(5)
Dep var:	Daily excess return on dollar-weighted calendar portfolio of herding events (%)				
Herding events:	All RH herding	All TAQ herding	RH herding only	TAQ herding only	Both RH and TAQ herding
$Alpha$	-0.552***	-0.328***	-0.479***	-0.168*	-0.722***
	(0.108)	(0.102)	(0.094)	(0.088)	(0.237)
Mkt_Rf	0.660***	0.506***	0.798***	0.580***	0.313
	(0.127)	(0.107)	(0.130)	(0.090)	(0.222)
SMB	0.446**	0.043	0.430**	-0.055	0.741
	(0.219)	(0.250)	(0.208)	(0.253)	(0.523)
HML	0.355	-0.238	0.312	-0.397	0.337
	(0.217)	(0.265)	(0.201)	(0.299)	(0.466)
MOM	-0.229	-0.675***	-0.182	-0.707***	-0.334***
	(0.166)	(0.234)	(0.162)	(0.270)	(0.293)
RMW	-0.968***	-1.279***	-0.437	-0.951***	-2.158
	(0.301)	(0.285)	(0.281)	(0.256)	(0.0562)
CMA	-0.773*	0.223	-0.766**	0.583	-0.733
	(0.424)	(0.431)	(0.362)	(0.355)	(0.990)
Observations	555	555	555	555	553
R-squared	0.256	0.163	0.353	0.205	0.054

Table A8: Summary Statistics for Robinhood Sales Herding Events

Sales herding events are securities in the bottom 0.5% of negative user change ratio on day t and a minimum of 100 users on day $t-1$. The variables are *users_close* (last observed user count for a stock prior to the 4 pm ET close), *users_last* (last user count of the day), *userchg* (daily change in *users_close*), *userratio* ($users\_close(t)/users\_close(t-1)$), *prc* (closing price), *size* (market cap in millions), *ret* (daily return), *openret* (overnight return), *dayret* (daytime return), *daily_buys* (number of TAQ daily retail buys), *daily_sells* (number of TAQ daily retail sells), *net_buys* ($daily\_buys - daily\_sells$), *taq_retimb* ($net\_buys/(daily\_buys + daily\_sells)$).

Variable	N	mean	sd	min	p25	p50	p75	max
users_close	4,889	986.70	4,321.67	0.00	163.00	330.00	860.00	147,351.00
users_last	4,889	982.20	4,311.89	0.00	162.00	329.00	852.00	148,547.00
userchg	4,889	-146.86	597.89	-19,643.00	-118.00	-43.00	-20.00	-5.00
userratio	4,889	0.87	0.09	0.00	0.86	0.90	0.92	0.96
prc	4,581	24.73	70.11	0.29	3.31	9.73	26.30	2,990.70
size ($mil)	4,178	1,169.30	3,533.03	0.11	24.79	189.55	924.23	104,742.80
ret (%)	4,581	-2.57	8.77	-66.03	-6.24	-1.60	1.69	69.81
openret (%)	4,572	-1.41	5.29	-72.54	-2.50	-0.45	0.51	64.47
dayret (%)	4,573	-1.17	7.42	-47.61	-4.65	-0.76	2.18	71.91
daily_buys	4,541	282.41	1,651.55	0.00	39.00	94.00	231.00	93,058.00
daily_sells	4,541	304.16	1,737.56	0.00	48.00	111.00	260.00	101,813.00
net_buys	4,541	-21.75	217.17	-8,755.00	-38.00	-11.00	5.00	5,821.00
taq_retimb	4,541	-0.09	0.20	-1.00	-0.19	-0.07	0.03	1.00

Table A9: The Persistence of Herding Events

The table summarizes the results of a linear probability model of *rh_herd* (*rh_negherd*) regressed on lags of *rh_herd* and *rh_negherd*. *rh_herd*(t) (*rh_negherd*(t)) is an indicator variable that takes a value of one if the percentage change in users is in the top (bottom) 0.5% for stocks with positive (negative) user changes on day t and a minimum of 100 users on day $t-1$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

	(1)	(2)
Dep var:	*rh_herd*(t)	*rh_negherd*(t)
rh_herd$(t-1)$	0.116***	0.192***
	(0.005)	(0.007)
rh_herd$(t-2)$	0.003	0.131***
	(0.002)	(0.006)
rh_herd$(t-3)$	0.010***	0.025***
	(0.002)	(0.004)
rh_herd$(t-4)$	0.008***	0.019***
	(0.002)	(0.003)
rh_herd$(t-5)$	0.002	0.009***
	(0.001)	(0.003)
rh_negherd$(t-1)$	0.004***	0.128***
	(0.002)	(0.006)
rh_negherd$(t-2)$	0.004**	0.045***
	(0.002)	(0.004)
rh_negherd$(t-3)$	0.006***	0.030***
	(0.002)	(0.004)
rh_negherd$(t-4)$	0.006***	0.016***
	(0.002)	(0.003)
rh_negherd$(t-5)$	0.006***	0.022***
	(0.002)	(0.003)
Observations	3,949,293	3,949,293
R-squared	0.014	0.097

Table A10: Regression of Daily Returns on Lagged Robinhood Buy/Sell Herding Indicator (Top/Bottom 0.5% Percentage User Change)

The dependent variable is the daily stock return (%) winsorized at the 0.1% level (*ret*). In columns (1) to (3), the key independent variable, $rh\_herd(t)$ ($rh\_negherd(t)$), is an indicator variable that takes a value of one if the percentage change in users is in the top (bottom) 0.5% for stocks with positive (negative) user changes on day t and a minimum of 100 users on day $t-1$. In columns (4) to (6) (or columns (7) to (9)), the $rh\_herd(t)$ ($rh\_negherd(t)$) indicator variable equals one if the prior conditions are met and the overnight return from the close on day $t-1$ to the open on day t is positive (negative). Control variables include retail order imbalance from TAQ (*taq_retimb*), lagged returns (*ret*), lags of an indicator variable if the *rh_herd* ($rh\_negherd(t)$) measure is missing (and *rh_herd* ($rh\_negherd(t)$) is set equal to zero), and day fixed effects. 5-day Buy (Sell) Herd AR (%) is the sum of the coefficients on the five lags of *rh_herd* (*rh_negherd*). Robust standard errors clustered by day are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep var:					$ret(t)$ (%)				
Herding Events:		All Events			Overnight Return > 0			Overnight Return < 0	
$rh\_negherd(t-1)$	-0.227**	-0.234*	-0.200	-0.234	-0.170	-0.171	-0.440***	-0.469***	-0.414**
	(0.112)	(0.123)	(0.125)	(0.159)	(0.159)	(0.161)	(0.155)	(0.178)	(0.179)
$rh\_negherd(t-2)$	-0.116	-0.010	0.002	0.337**	0.357**	0.382***	-0.515***	-0.336**	-0.334**
	(0.106)	(0.115)	(0.116)	(0.145)	(0.142)	(0.144)	(0.142)	(0.154)	(0.155)
$rh\_negherd(t-3)$	-0.081	-0.069	-0.075	-0.247*	-0.207	-0.205	-0.083	-0.089	-0.090
	(0.100)	(0.107)	(0.106)	(0.137)	(0.137)	(0.139)	(0.138)	(0.151)	(0.151)
$rh\_negherd(t-4)$	-0.087	-0.125	-0.127	0.023	0.019	0.024	-0.222*	-0.292**	-0.289**
	(0.101)	(0.100)	(0.102)	(0.133)	(0.133)	(0.135)	(0.134)	(0.140)	(0.142)
$rh\_negherd(t-5)$	-0.121	-0.161*	-0.150	-0.012	-0.023	-0.016	-0.212*	-0.268**	-0.253*
	(0.084)	(0.093)	(0.094)	(0.115)	(0.115)	(0.117)	(0.116)	(0.129)	(0.130)
$rh\_herd(t-1)$	-1.340***	-1.042***	-1.112***	-1.756***	-1.133***	-1.276***	-0.515**	-0.972***	-0.896***
	(0.160)	(0.171)	(0.175)	(0.215)	(0.257)	(0.260)	(0.209)	(0.227)	(0.229)
$rh\_herd(t-2)$	-0.717***	-0.767***	-0.759***	-1.134***	-1.223***	-1.200***	-0.126	-0.111	-0.126
	(0.130)	(0.139)	(0.139)	(0.176)	(0.221)	(0.220)	(0.175)	(0.204)	(0.204)
$rh\_herd(t-3)$	-0.242*	-0.152	-0.153	-0.387**	-0.161	-0.142	-0.204	-0.371*	-0.394*
	(0.125)	(0.132)	(0.130)	(0.163)	(0.204)	(0.200)	(0.195)	(0.211)	(0.211)
$rh\_herd(t-4)$	-0.275**	-0.242*	-0.244*	-0.310**	-0.173	-0.184	-0.330**	-0.510***	-0.489***
	(0.117)	(0.129)	(0.130)	(0.145)	(0.185)	(0.184)	(0.163)	(0.183)	(0.184)
$rh\_herd(t-5)$	-0.185	-0.240*	-0.228*	-0.294*	-0.338*	-0.322*	-0.122	-0.195	-0.190
	(0.122)	(0.133)	(0.133)	(0.152)	(0.192)	(0.192)	(0.167)	(0.173)	(0.174)
Observations	3,656,926	3,652,401	3,312,553	3,656,926	3,652,401	3,312,553	3,656,926	3,652,401	3,312,553
R-squared	0.196	0.199	0.205	0.196	0.199	0.205	0.196	0.198	0.205
Days	550	550	550	550	550	550	550	550	550
Lags of Retail OI	NO	YES	YES	NO	YES	YES	NO	YES	YES
Lags of Return	NO	NO	YES	NO	NO	YES	NO	NO	YES
5-day Sell Herd AR (%)	-0.632	-0.600	-0.550	-0.131	-0.0251	0.0145	-1.472	-1.454	-1.380
Neg. Std. Err.	(0.213)	(0.211)	(0.215)	0.294	0.290	0.297	0.306	0.314	0.318
5-day Buy Herd AR (%)	-2.759	-2.444	-2.496	-3.881	-3.030	-3.124	-1.296	-2.160	-2.095
Pos. Std. Err.	(0.330)	(0.353)	(0.351)	0.441	0.528	0.525	0.425	0.453	0.456