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

November 2020

We appreciate the comments of Ivo Welch.

Attention Induced Trading and Returns: Evidence from Robinhood Users

Abstract

Consistent with attention-induced trading models' predictions, we link episodes of intense buying by retail investors at the brokerage Robinhood to future negative returns. Average five-day abnormal returns are -3% (-6%) for the top stocks purchased each day (more extreme herding) by Robinhood users. We find that herding episodes are related to the simplified display of information on the Robinhood app and to established proxies for investor attention. These factors lead to more concentrated trading by Robinhood users that can impact pricing. For example, during Robinhood outages, retail investor volume drops significantly among stocks that are likely to capture investor attention.

During the last half a century, the biases and heuristics that influence investor decisions have not changed. However, the environment in which these decisions are made has changed dramatically. For individual investors, two of the most consequential changes are an explosion in the types and of sources information and a virtual elimination of frictions associated with trading.

In 1971, most individual investors obtained their information by watching the evening news on ABC, NBC, or CBS, watching PBS's Wall Street Week with Louis Rukeyser on Friday nights, reading the financial pages of their local newspaper or investment newsletters that arrived by mail, or by calling their broker for recommendations. Today investors can watch CNBC 24/7, read the Wall Street Journal online, visit thousands of websites devoted to financial markets, review analyst forecasts, examine Securities and Exchange Commission (SEC) filings, and receive investment alerts on their mobile phones.

The changes to how individual investors trade are equally as dramatic. In 1971, an investor had to call her broker during business hours to place a trade. She paid a minimum commission of $160.50 to purchase 100 shares of a $10 stock (Jones, 2000). And trading costs were higher for odd lots. Today, she can buy a fractional share, commission-free, day or night with a few clicks on her smartphone.

A pioneer of some of these trading changes is the brokerage Robinhood. Robinhood was the first brokerage to offer commission free trading. Robinhood's app is simple and engaging, designed to encourage people to invest. 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] Robinhood also innovated how information is presented to investors. Specifically, Robinhood provides less detailed stock level information than many other brokerage firms. TD Ameritrade, for example, provides over 400 charting indicators for each stock, while Robinhood provides five.[2] Instead, Robinhood prominently displays a simple intuitive form of information: stock lists, such as "Top Movers" and "100 Most Popular."

Barber and Odean (2008) argue that limited attention prevents retail investors from considering all available information and possible stock choices. Instead, many retail investors choose stocks to buy from the subset of stocks that catch their attention. Because most investors own only a few stocks and do not sell short, limited attention plays a smaller role in their sales decision. This leads retail investors to be strongly on the buy side of the market for stocks that attract a lot of attention.

When buying stocks, investors with accounts at Robinhood (Robinhood users) are likely to be more influenced, both individually and as a group, by limited attention than other investors for several reasons. First, half of Robinhood users are first time investors[3] who are unlikely to have developed their own clear

[1] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html.
[2] https://www.stockbrokers.com/compare/robinhood-vs-tdameritrade
[3] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html.

criteria for buying a stock. Inexperienced stock investors are likely to be more heavily influenced by attention (Seasholes and Wu, 2007) and by biases that lead to return chasing (Greenwood and Nagel, 2009). Second, the Robinhood app directs Robinhood user's attention to the same small subset of stocks, such as the 20 "Top Movers," while offering limited additional information that might lead to more heterogeneous choices. 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 (System 1 in Kahneman (2011)) and less on critical thinking (System 2). Fourth, Robinhood users may deliberate and hesitate less than other investors when trading due to a lack of frictions; it is very easy to place trades on the Robinhood app and the ostensible cost of trading—i.e., commissions—is zero. Fifth, as evidenced by turnover rates many times higher than at other brokerage firms, Robinhood users are more likely to be trading speculatively and less likely to be trading for reasons such investing their retirement savings, liquidity demands, tax-loss selling, and rebalancing. The lack of non-speculative trading motives increases the potential for attention driven trading. Because 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 herd more than other investors.

In this paper, we find that Robinhood users are more subject to attention biases and more likely to chase stocks with extreme performance and volume than other retail investors. We systematically identify the Robinhood herding episodes and document that these episodes are followed by abnormal negative returns. We show that Robinhood herding is influenced by information that is prominently displayed on the Robinhood app. And we show that Robinhood herding can be forecasted by attention measures, such as lagged absolute returns and lagged abnormal volume, previously show to affect the by buy-sell imbalances of retail investors.

Our primary empirical analysis examines abnormal returns following events in which the number of Robinhood users owning a particular stock increases dramatically in one day. To preview the results, Figures 1 graphs buy and hold abnormal returns (BHARs) and Robinhood user changes around our identified herding events. Panel A defines herding events as the top 0.5% of positive user changes as a percent of prior day user count each day. Panel B defines extreme herding events as a user increase of more than 1,000 and more than 50% relative to the previous day. Both Figures graph a 31-day period from 10 trading days before the event day to 20 trading days after. The return and user patterns are similar. The average abnormal return on the herding day is 14% (42%). However, over the subsequent month, we observe a negative return of almost 5% (9%). In summary, large increases in Robinhood users are often accompanied by large price spikes and are followed by reliably negative returns. While some users profit from these episodes, we find that, in aggregate, Robinhood users who establish new positions during these episodes incur losses.

Herding by a few investors is unlikely to move prices in all but the least liquid stocks. There are, however, more than a few Robinhood users. Robinhood users grew from one million in 2016 to 13 million in May 2020, more users than Schwab (12.7 million) or E-Trade (5.5 million) had at the end of 2019. Additionally, 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."[4] Indeed, we show that during Robinhood outage events that retail participation in the top Robinhood stocks dropped by a statistically and economically meaningful amount. Furthermore, it is likely that the purchases of other retail investors are positively correlated with those of Robinhood users.

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). We also observe negative returns following a day when we observe both a surge in Robinhood users and the stock's price goes down. Aggressive retail buying in response to sharp price drops is consistent with Barber and Odean's (2008) attention theory, but a price drop following a price drop is not a reversal. One possible reason why stock prices drop after a negative price move accompanied by aggressive Robinhood user buying is that this buying slowed the stock's response to negative news. 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.

There are, however, several reasons to believe that post-herding negative returns we observe are caused, at least in part, by the trading of Robinhood users and other retail investors. Trading by small investors is more likely to influence the returns of smaller cap stocks. Consistent with this logic, 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. Retail trading has increased significantly at Robinhood and elsewhere in the post-Covid period (i.e, after March 13, 2020) and the negative return effect following Robinhood herding events is more pronounced in the post-Covid period. Note that our return analysis is primarily focused on Robinhood herding events, not on the long-term aggregate performance of Robinhood users, a topic addressed in Welch (2020).

We believe that design of the Robinhood app helps create the herding episodes we study in two ways. First, the app prominently displays lists of stocks in an environment relatively free of complex information. Frydman and Wang (2020) show that what information is displayed prominently on a brokerage's online trading screen influences investor behavior. Kaniel and Parham (2017) show that whether mutual funds do

[4] https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html

or do not appear on prominently displayed lists affects mutual fund flows.[5] Second, the Robinhood app makes it very easy to place trades. Barber and Odean (2002) show investors who access easier trading by switching from phone-based to online trading, trade more and perform worse for up to two years after switching.

The two most prominently displayed lists on the Robinhood app are the “Most Popular” list and the “Top Movers” list. “Most Popular” lists the 100 stocks with owned by the most Robinhood users. The list is long and largely static. It is static because it ranks stocks on the level of Robinhood ownership which changes very little from day to day. The top stocks on the Most Popular list tend to be large firms such as Apple, Tesla, and Microsoft and stocks such as Ford and GE which meet the criteria of Robinhood’s free share program.[6] These stocks consistently experience positive increases in users over time, but it’s difficult to disentangle whether the observed increase is a result of the “Most Popular” list or new users joining Robinhood with similar preferences to existing users.

In contrast, “Top Movers” lists only 20 stocks and changes every day (and throughout each day). This list is based on the absolute percentage price changes from the previous day close. Listing both positive and negative top movers in order on the same list is rare. There are many sites on the Internet where one can view the stocks with the largest daily gains, e.g., Yahoo! Finance Day Gainers[7], and sites with the largest daily losses, e.g., Yahoo! Finance Day Losers. Thus retail investors can find lists of top movers outside Robinhood. However, these lists are separated into gainers and losers rather than combined; and top gainers tend to be more prominently displayed. Additionally, 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.

If, as we claim, the buying of Robinhood users is more attention driven, we would expect Robinhood users to be buying stocks on Robinhood’s Top Mover list more aggressively than other retail investors. Indeed, this is the case.

To demonstrate that the app itself is driving Robinhood traders’ trading, we look at large positive and negative movers separately. Since the app lists top movers – whether positive or negative in order – Robinhood users’ attention is being directed to stocks with large real time gains and losses. However, non-Robinhood retail investors likely focus more on real time gains than losses. Thus, if the app drives trading, we would expect Robinhood users to heavily buy both gainers and losers, while other retail investors will have more of a tendency to buy gainers. This is precisely what we find: Robinhood users are drawn to

[5] Da et. al. (2018) find that recommendations of an advisory firm followed by many Chilean pension investors generate correlated fund flows and influence market returns.
[6] A notable exception is Aurora Cannibas Inc (ACB), which was the most popular Robinhood stock during much of the sample period until it engaged in a reverse split that caused its rank to drop from 1 to 7.
[7] While some websites refer to stocks with the largest daily returns as “Top Winners,” the term “Top Gainers” appears to be more popular.

trading both extreme gainers and losers while other retail investors prefer to buy extreme gainers rather than losers.

As noted previously, during our sample period, a strategy of selling after a Robinhood herding event and repurchasing five days later would have resulted in a BHAR of 3.5% (6.4% for extreme herding events). For the 4,884 herding events we observe, this strategy would have yielded a positive BHAR 63% percent of the time. One would expect astute investors to have exploited such profitable and reliable opportunities. And, indeed, it appears that some did.

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. And we do find a marked increase in short selling for stocks involved in Robinhood herding events. We rank stocks on the maximum daily return during a two-week period and then separate them into those with and without Robinhood herding events during the period. Stocks with Robinhood herding events show much larger percent increases in short interest than similarly ranked stocks without Robinhood herding. Specifically, for the stocks with the top 25 returns for the period, the average change in short interest is three times greater for stocks that experienced Robinhood herding events.

Our study is of particular interest given the unique dataset of the retail investors (Robinhood users) that we analyze. To our knowledge, three 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.[9] In a JP Morgan report, Cheng, Murpy, & Kolanovic (2020) show 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 retail trading generally lowered stock bid-ask spreads and price impact of trades. In contrast, we find that stocks with unusually high levels of interest among Robinhood users experience negative returns.

In summary, we provide two main contributions to the academic literature. First, consistent with the predictions of attention-induced trading models, we link episodes of intense buying by retail investors to negative returns following the herding episodes. This result contributes to the literature that documents price reversals following attention grabbing events: Jim Kramer's stock recommendations (Engelberg,

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

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, et al., 2020), and repeat news stories (Tetlock, 2011). While most of this literature has documented negative returns subsequent to positive return events, we also find negative returns following negative return events. Rather than identifying events using proxies for attention, we analyze the intense buying episodes by Robinhood investors, who we argue are more likely to engage in attention motivated trading than other retail investors. We then show these episodes are related to factors (e.g., extreme returns, unusual volume, and information display) that likely capture the attention of investors. Moreover, when Robinhood experiences outages, we observe the largest decrease in retail trading among stocks that attract the attention of Robinhood users (the most popular stocks on Robinhood and stocks with a high probability of a herding event).

Our second contribution is to present evidence that the display of information 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) returns, and changes regularly, has a large effect on Robinhood users. Importantly, as discussed above, the sorting of stocks on absolute return is different from the most prevalent stock rankings which highlight stocks with the biggest gains. Consistent with the idea that this prominent list influences the behavior of Robinhood users more than other retail investors, we show that Robinhood users are more likely to buy stocks with either extreme negative or positive returns; in contrast, the general population of retail investors is more likely to buy stocks with extreme positive returns rather than stocks with extreme negative returns.

This result fits into the emerging literature that emphasizes the display of information can affect 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). Displaying return performance for index funds can lead investors to prefer high-fee funds (Choi, Laibson & Madrian, 2009) and prominently featuring expense information can lead investors to prefer low-fee funds (Kronlund, Pool, Sialm & Stefanescu, 2020). 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.

1 Data and Methods

In this section, we describe the main Robintrack dataset which 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.[10] 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).[11] Our main results include all Robintrack securities, since we do not have strong priors about what types of securities will experience herding events.[12]

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 (0.006,0.01); retail sells are trades reported to a FINRA TRF with fraction penny prices between (0.00,0.004).[13]

In Figure 2, 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 reported having 13 million users, which translates into about 3 stock positions per user.[14] The red line in the figure denotes the date on which the COVID national emergency was declared in the US (March 13, 2020); there is a clear increase in Robinhood users after this date. In Figure 2, Panel B, we plot the total

---

[10] There are about 11 dates during the sample period that are missing user data, four in Jan 2019 and seven in Jan 2020. There are 16 dates where we observe users, but there are no observations between 2 and 4 pm ET.

[11] 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.

[12] 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%.

[13] 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.

[14] 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)

number of TAQ retail trades for comparison. There is also an increase in retail trading 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 2, Panel C, we plot the sum of the absolute value of Robinhood user changes (green line) and the absolute value of TAQ net buying (blue line). Both 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 key variables in our later analysis of herding events are $users\_close$, which 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 analysis 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 are likely due to the fact that 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. We employ two strategies to catalog these herding events. In our first strategy, 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).

In the second strategy, we employ a more stringent filter. We identify stocks with an increase of at least 1,000 users ($userchg(t) > 1{,}000$) and at least a 50% increase in users ($userratio(t) \geq 1.5$). This more stringent definition results in the identification of 900 herding events, which we denote with an indicator variable rh_herd2. The majority of these events occur during the pandemic period (610 or 58%), which is to be expected given the large increase in Robinhood users during this period.

In Table 2, we present descriptive statistics on the herding events across stock-day observations. In Panel A, we focus on the *rh_herd* events. 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 1/3rd) 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.

In Panel B, we focus on the more extreme *rh_herd2* events. The user changes and returns around the episodes for these stocks are more dramatic than those observed for the first definition, which is what we expected since these are more extreme herding events.

2 Attention and Stock Selection

In the first part of our analysis, we document that Robinhood users show excessively concentrated trading activities, compared to the general population of retail investors measured by TAQ dataset. To establish these concentrated activities are induced by attention, we then explore a few unique features of the Robinhood platform and how these features affect the trading activities of Robinhood users. In addition, we assess the economic significance of Robinhood trading, especially the attention-induced trading, by analyzing the change in retail volume on TAQ when Robinhood outages occur.

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 that 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 (measured by user changes). We separately identify the 10 stocks with the most TAQ retail buying (measured by net buying). Among stocks with positive net buying, we then calculate the percentage of total net buying that is observed in each of the top 10 stocks for Robinhood and TAQ retail investors. 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 calculate for net selling.

In Figure 3, we present the mean concentration of buying (Panel A) and selling (Panel B) for Robinhood users (green bars) and TAQ retail trades (purple 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.

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 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.

2.2 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.[15] Presented with a large variety of stimuli, investors using traditional investing products are likely to have heterogenous responses given the limited capacity of human attention and the allocation of attention is highly flexible (Kahneman, 1973). In contrast, the reduced number of stimuli on Robinhood makes it easy for investor to focus their attention and likely generate coordinated attention-induced responses.

[15] https://www.stockbrokers.com/compare/robinhood-vs-tdameritrade

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." As discussed above, the two most prominently displayed lists the "Most Popular" list and the "Top Movers" list. We focus on the "Top Movers," list because it is short and constantly changing while the "Most Popular" list is long and largely static. "Top Movers" lists stocks with the day's largest percent gains and losses since the market close from the previous day. The default sorting is based on the absolute returns and thus mixes top gainers and top losers.[16] 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 separately report the top gainers and top losers rather than mixing them together.

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 salience for the two groups of stocks is similar. 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 salient on other media accounts when they are reported separately from the top gainers. Accordingly, we would expect general retail investors respond less strongly to top losers than top gainers.

Figure 4 presents the graphical evidence for the comparison. We measure the buying activity of Robinhood investors by intraday user change[17] 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). In Panel A, stocks are ranked 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. Thus buying activity is measured subsequent to ranking. Th 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; 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). [18]

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

[16] Figure A1 provides an example of the "top mover" list on Oct 8th, 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 in this example.

[17] 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.

[18] Both RH intraday user change and TAQ intraday net retail buying are adjusted for day fixed effects.

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:

$$NetBuy_{it} = \beta_0 + \beta_1 Score_{it} + \beta_2 I_{R_{it}<0} + \beta_3 Score_{it} \times I_{R_{it}<0} + \alpha_t + \varepsilon_{it}, \tag{1}$$

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.

In Table 4, Columns (1) and (2) sort top mover scores on absolute overnight returns. The constant term is statistically significant and positive, suggesting both Robinhood investors (Column (1)) and general retail investors (Column (2)) are net buyers of these top mover stocks. Moreover, the buying activity increases with top mover scores in both columns. This indicates that 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 228.5 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 eight ($\approx 228.5/30.16$). 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. Columns (3) and (4) sort top mover scores on absolute daily returns and find similar results.

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

2.3 Attention Proxies and Robinhood Herding Events

Previous sections show that Robinhood investor's attention and buying activity is influenced by the salient lists compiled by the platform (e.g., "Top Movers", "100 Most Popular", etc.), which are likely one driver for the herding episodes identified in Section 1. Here, we examine the relation between a collective set of attention measures and the herding episodes we study. To do so, we estimate a linear probability model by regressing the extreme herding episode indicator on a set of attention measures, including absolute lagged returns, lagged user change, lagged level of users and lagged abnormal volume. 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 5 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 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 platform and leads 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 10, the probability of this stock being heavily bought the next day increases by 6.7% (Column (1)); this is statistically significant at 1% level. The attention-grabbing effect of extreme absolute returns is more generally captured by the positive and significant coefficient on the level of the absolute returns (Column (2)). Column (3) separately considers positive and negative returns. As the result shows, negative returns are in general less likely to generate extreme herding of Robinhood investors. The insignificant interaction term suggests that investors respond more strongly to larger price swings for both positive and negative returns. In addition, we also look at attention measures such as lagged user change, lagged level of users and lagged abnormal volume, which 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.

2.4 The Effect of Robinhood Outages on Retail Trading Volume

To establish the importance of Robinhood user trading, 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, but particularly on stocks that are popular among Robinhood users or stocks that are good candidates for the herding events we study.

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 that trade using other brokers. 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). Retail trades are identified in TAQ as in Boehmer et al. (BJJZ 2020).

In Figure 5, 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 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:

$$RetailPerc_{it} = a + bOutage_t + \mu_{tod} + \mu_{stock} + e_{it}\,, \tag{2}$$

where $RetailPerc_{it}$ is the percent of trades that are retail trades on TAQ during period *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:

$$RetailPerc_{it} = a + bOutage_t + cRepair_t + \mu_t + \mu_{tod} + \mu_{stock} + e_{it}\,. \tag{3}$$

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:

$$RetailPerc_{it} = a + bOutage_t + cPartial + dRepair_t + \mu_t + \mu_{tod} + \mu_{stock} + e_{it}\,. \tag{4}$$

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 6 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 (1), Table 6). 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 (12.1%) during this period. For the 50 most popular Robinhood stocks, retail trading declines by 5.22 ppt ($p < .001$), which represents 35% of the average fraction of retail trading (14.6%) for these stocks. For the 50 high attention stocks, we observe a 4.37 ppt decline in trading volume, which represents a 26% decline in the typical level of retail trading volume for these stocks (16.58%).

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 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 trading volume in stocks (and 1/3rd of retail trading volume) in the 50 most popular Robinhood stocks. Perhaps more 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 about 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.

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 7, 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:

$$BHAR_{i\tau} = \prod_{t=\tau-10}^{\tau}(1 + R_{it}) - \prod_{t=\tau-10}^{\tau}(1 + R_{mt}) \,. \qquad (5)$$

[22] In Figure A2, we show the five-minute mean trading during the 11:35 am to 12:40 pm time interval. There is a noticeable increase in trading volume 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.

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.[23]

We report results for *rh_herd* in Panel A (top 0.5% of daily buys) and *rh_herd2* in Panel B (more extreme episodes). In both cases the pattern is similar. 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% and 42% for the rh_herd and rh_herd2 respectively. Interestingly, many stocks have negative returns the day prior to and 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, although again similar for both groups. Immediately after the herding event returns turn significantly negative. After just five days, stocks in the Top 0.5% experience negative abnormal returns of -3.5% whereas the more extreme group experiences a negative return of over 6%. By the end of the 20-day period, the return decline totals almost 5% (9%) for the rh_herd (rh_herd2) group. These results are economically and statistically significant. These results are not driven by just a few stocks. In both cases, 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 1, we plot the buy-and-hold abnormal returns for our rh_herd and rh_herd2 events. Results over the entire period are reported in Panel A whereas *BHAR*s starting at Day 1 are reported in Panel B. The pattern of returns around herding events is quite clear.

[23] 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).

Robinhood users are attracted by extreme return events. Their coordinated buying leads to price pressure and then a subsequent poor return performance.

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 the follows the herding episodes, we construct a calendar-time portfolio. Define P_{it} as the closing price for stock *i* on day *t*. The portfolio invests in $\$1/P_{it}$ shares of each event stock at the close of trading on the day of the herding episode and holds the shares for five days. The day t portfolio return is calculated as:

$$R_{pt} = \sum_{i=1}^{N} w_{it} R_{it}, \tag{6}$$

where

$$w_{it} = \frac{S_{i,t-1} P_{it-1}}{\sum S_{i,t-1} P_{it-1}}. \tag{7}$$

R_{it} is the stock return, w_{it} is the weight, and S_{it} is the total number of shares held for stock *i* on day *t*:

$$S_{it} = \sum_{h=-4}^{0} \frac{I_{i,t-h}}{P_{i,t-h}}. \tag{8}$$

The numerator is an indicator variable, $I_{i,t-h}$, that equals one if day *t-h* is an episode day for stock *i*. Thus, shares are purchased on herding event days (but not on other days) at the closing price for the day. Some stocks will have multiple events within a five-day window (and hence the summation of shares across days *t*-4 to *t*.

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:

$$R_{pt} - R_{ft} = \alpha + \beta\left(R_{mt} - R_{ft}\right) + \sum_{k=1}^{K} c_k F_t^k + e_{pt} \tag{9}$$

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 8 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 -54 to -60 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 -79 to -94 bps per day.

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:

$$R_{it} = a + \sum_{k=1}^{5} b_k \, rh\_herd_{i,t-k} + \sum_{k=1}^{5} c_k \, taq\_retimb_{i,t-k} + \sum_{k=1}^{5} d_k \, R_{i,t-k} + \mu_t + e_{it} \,. \qquad (10)$$

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 plagues 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; Lehman, 1990; Lo & Mackinlay, 1988, 1990; Campbell, Grossman & Wang, 1993), though there has been a large decrease in the return reversal strategy after 2000 (Khandani & Lo, 2011). These early analyses speculate that the origins of return reversals may emanate from overreaction, the rewards to providing liquidity to markets with shortage of buyers/sells, 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. Following this logic, the inclusion of lagged returns may be overcontrolling as both return reversals the reversals we document may have similar origins in attention motivated trading.

Table 9 presents results for our herding measure based on the top 0.5% of daily Robinhood user changes. We present results separately for overnight returns (Columns (1)-(3)), daytime returns (Columns (4)-(6)), and daily returns (Columns (7)-(9)). There is weak evidence of anomalous behavior in overnight returns, but consistently negative abnormal returns in the five days after the herding event. The introduction of lagged retail order imbalance and lagged returns has little impact on these results. 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% and all of these losses can be traced to intraday day returns.

Table 10 presents results for the more extreme herding measure (at least a 50% increase in Robinhood users representing a minimum of 1000 new users). The results are generally consistent, but more dramatic. In the last row of Columns (7)-(9), we observe a five-day abnormal return of 5.4%.

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

3.4 Subsample Analyses

To test the robustness of these patterns, we conduct a battery of robustness tests. In our first analysis, we only include Robinhood herding events that are accompanied by negative returns on the same day (i.e., a big increase in Robinhood users despite a negative same-day return), which represent about 1/3rd of all events. Table 11, Panel B, shows we observe similar return patterns for this subsample and the observed magnitudes are only slight smaller than those observed for the full sample. The negative return sample allows us to rule out the broader empirical phenomenon of return reversals as the origin of the observed return pattern. By focusing on only events with same-day negative returns, we provide novel evidence that the buying behavior of attention motivated traders are providing price support when prices are falling.

Because Robinhood users can trade with limited capital, low prices stocks are appealing to them and we thus include them in our analysis. It's natural to wonder whether our results are driven by us being more likely to observe closing prices at bid prices on the day of the herding event and thus generates observations of negative returns on the next day, which is more likely to have a closing price at the bid and ask. 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 C). We also find reliably negative returns for stocks with prices in excess of $5, but the return magnitudes are smaller (Panel D).

We anticipate the negative returns we document will be present in small cap stocks, but small 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 dollar 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 E and F).

Our main results include both stocks and other securities (e.g., ADRs and ETFs). We observe similar return patterns for these securities (Panel G).

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 H and I).

3.5 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 much less dramatic user changes 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 Table A2). 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 Table A3), 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 9. 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 Table A4), perhaps because they accelerate the poor performance that follow purchase herding events. In unreported results we focus on sales herding events that do not follow a purchase herding event in the prior five days and do not find evidence of negative return drift.

3.6 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 excess 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 $\Delta u_{i\tau}$ 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:

$$PPrc_j = \frac{\sum_{\tau=-10}^{+20} \mathrm{UI}_{\mathrm{i\tau}} \Delta u_{i\tau} P_{i\tau}}{\sum_{\tau=-10}^{+20} \mathrm{UI}_{\mathrm{i\tau}} \Delta u_{i\tau}}. \quad (11)$$

The indicator $UI_{i\tau}$ equals one on days when the change in users is positive, $\Delta u_{i\tau} > 0$.

We then compute the analogous calculation for average sales price on days where 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 (i.e., $u_{i,+20} = 0$):

$$SPrc_j = \frac{\sum_{\tau=-10}^{+20} (1 - \mathrm{UI}_{\mathrm{i\tau}}) \Delta u_{i\tau} P_{i\tau}}{\sum_{\tau=-10}^{+20} (1 - \mathrm{UI}_{\mathrm{i\tau}}) \Delta u_{i\tau}}. \quad (12)$$

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

$$PrcRatio_j = \frac{SPrc_j}{PPrc_j} - 1 \quad (13)$$

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 where 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 6.

On average, the Robinhood community loses approximately 5% during each herding event using raw returns. After adjusting for the market return, losses are slightly more than 6%. Both results are highly significant both statistically and economically. There are slight differences between the pre-COVID and COVID periods, although both suggest similar outcomes. Overall, these findings suggest that extreme herding cause negative wealth outcomes for the overall Robinhood community.[24]

4 Short Interest and Herding Events

In the prior section, we find that Robinhood users' herding can lead to price pressure on those stocks and then 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 participates could trade against Robinhood users' order flow in these herding situations in an attempt to profit off Robinhood users'

[24] In untabulated results, we include users that hold the stock at the beginning of the herding period. If we include them, the market adjusted returns are -5.5%. In Figure A3, we show that in the average event about 2/3rds 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 14.6 million to 3.5 million.

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 these disclosures open themselves up to front-running and as well as allowing others to profit off their private information.[25] Indeed, researchers have documented these disclosures are used by market participants (e.g. Brown & Schwarz, 2020). Thus, it is somewhat surprising Robinhood would voluntarily make these data available to the general public.[26]

In this section, we evaluate whether other market participants attempt to profit from the predictably negative returns that follow Robinhood herding events. To do so, we collect short interest data from the NASDAQ website. The NASDAQ posts bi-monthly short interest data for all its stocks for the last year. Thus, we have data on all NASDAQ stocks from September 2019 to August 2020. We compute the change in short interest each approximately two-week period as the change in short interest divided by the prior level. We require the prior level of short interest to be at least 100 shares. We also winsorize the top 0.1% of to control for extreme outliers.

To begin our analysis, we examine the relation between top stock performance, Robinhood herding, and short interest. Specifically, during the two-week period between short interest observations, we calculate the highest daily return rank for each stock since we expect short interest changes 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. We then calculate the average change in short interest for each return rank for the non-Robinhood herding and Robinhood herding groups separately. In Figure 7, 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, changes in short interest are indeed higher for stocks that had a high daily return rank during the two period. Second, we see in all cases the short interest changes 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 short interest changes on either the maximum daily return (*Max_ret*) or the maximum return ranking (*Max_rank*) during the two week period. Our primary variables of interest are either the average user change ratio during the

[25] It is not clear 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 et al. (2013) find that after controlling for biases that hedge fund returns themselves have little alpha.
[26] Robinhood ceased releasing user data in August 2020 stating 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.)

period (*rh_chgratio*) or whether the stock was herded by Robinhood users (*rh_herd*). We average across periods and obtain standard errors and t-values via Fama-Macbeth (1973).

We find results consistent with the previously reported figure. A 100% increase in (doubling) the average change ratio leads to at least a 157% increase in short interest (Column (1)), which is both statistically and economically significant. If a stock is herded by Robinhood users (Columns (2) and (4)), short interest is more than 100% 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 experience negative average returns of approximately 5% over the next month. More extreme herding events are followed by negative average returns of approximately 9%.

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 the 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), 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., & 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. (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., Charles M. J., Zhang, X., & Zhang, X. (2020). Tracking Retail Investor Activity. *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.

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.

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

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

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.

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.

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.

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.

Jones, C. M. (2002). A Century of Stock Market Liquidity and Trading Costs. *Working Paper*.

Kaniel, R., Liu, S., Saar, G., & Titman, S. (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. *Journal of Finance*, *67*(2), 639-680.

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.

Khandani, A. E., & Lo, A. W. (2011). What Happened to the Quants in August 2007? Evidence from Factors and Transactions Data. *Journal of Financial Markets*, *14*(1), 1-46.

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*.

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

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

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*.

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

Ozik, Gideon and Sadka, Ronnie and Shen, Siyi, Flattening the Illiquidity Curve: Retail Trading During the COVID-19 Lockdown (November 10, 2020). Available at SSRN: https://ssrn.com/abstract=3663970

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*.

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

Appendix

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

Another way to examine how investors perform during herding episodes, we look at the number of users that 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 ($N^{buy}\_Pos_j$; $N^{buy}\_Neg_j$).

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

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

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:

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

The calculations above ignore periods when we experience a decrease in users. For these periods, we conduct a similar analysis where 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 $PercPos_j^{sell}$:

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

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

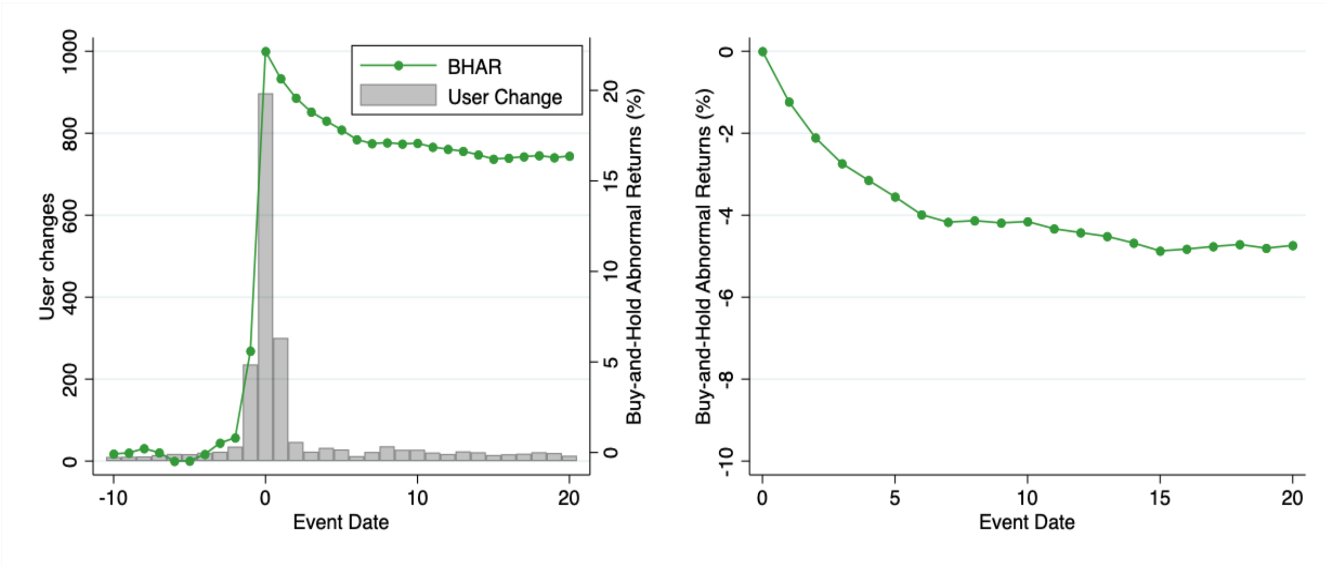
$$PercPos_j^{sell} = \frac{N^{sell}\_Pos_j}{(N^{sell}\_Pos_j + N^{sell}\_Neg_j)}. \quad \text{(A.6)}$$

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 Figure A3.

For buys, we find results consistent with the previously reported investor experiences. Only 38% 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 Robinhood users breakeven, the number of buy events (approximately 15 million) far exceeds sell events (approximately 3.5 million). The weighted average win rate is 42%.[27]

[27] The reason the number of sell events is so low is 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.

Panel A: Herding Events (Top 0.5% Percentage User Change)



Panel B: Extreme Herding Events (At Least 50% Increase in Users and 1000 New Users)

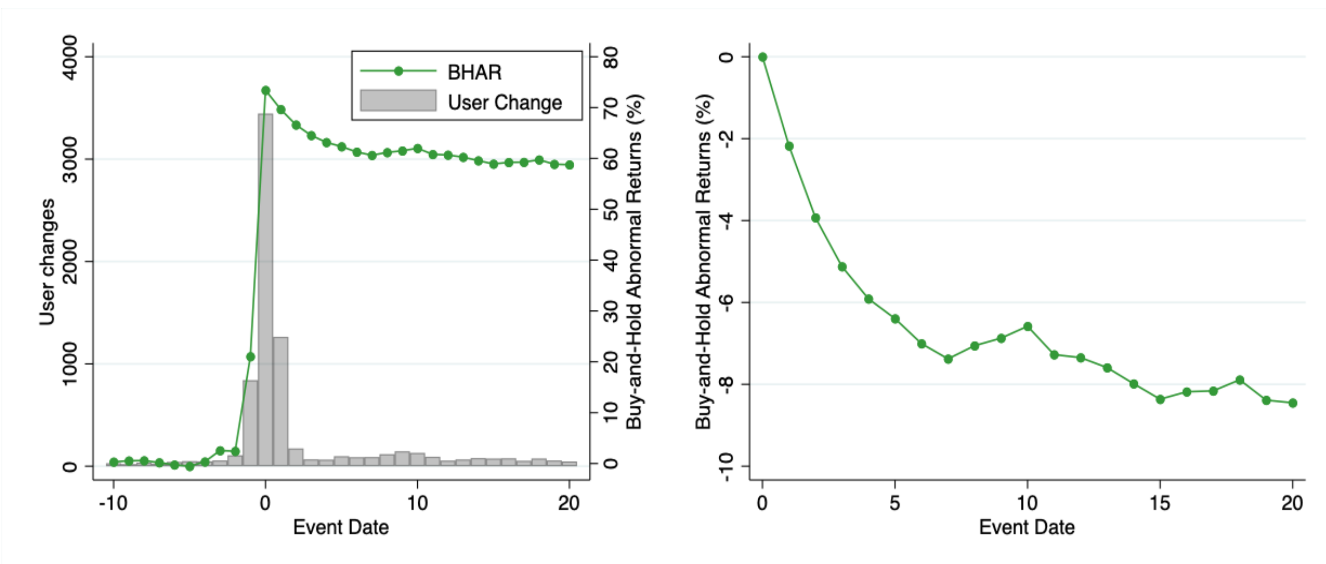
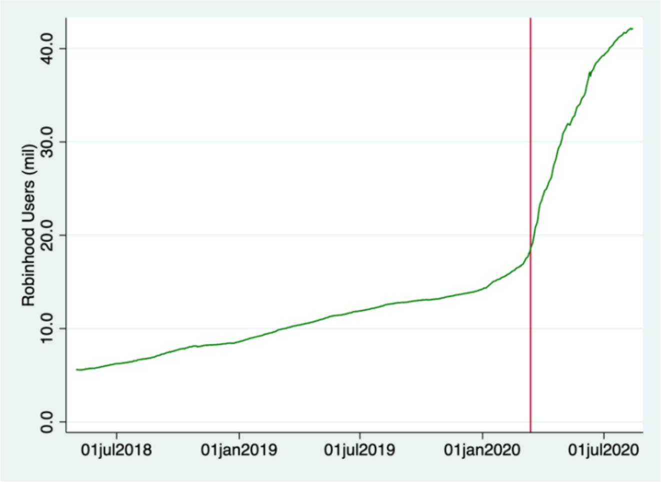


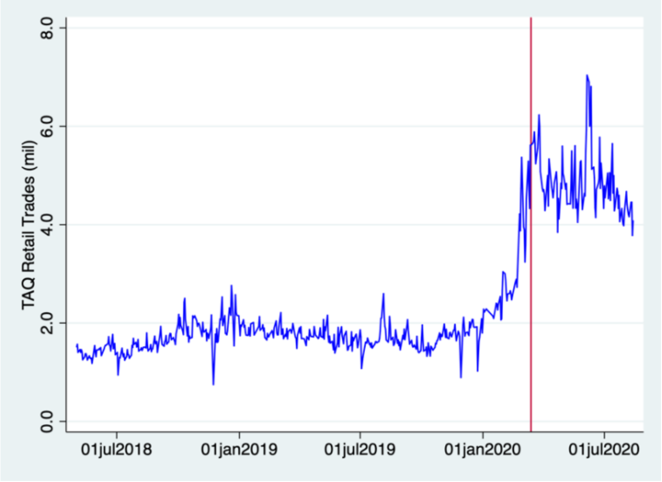
Fig. 1. Returns around Herding Events

The figures on the left depict 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 figures on the right display post-event mean BHARs starting from day 0. In Panel A, herding events are defined as the top 0.5% of user changes as a percent of the prior day user count. In Panel B, an extreme herding event occurs when the number of users increase by at least 1000 in one day and this increase is 50% or more of the prior day user count.

Panel A. Robinhood Total User Stock Holdings



Panel B. TAQ Total Retail Trades



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

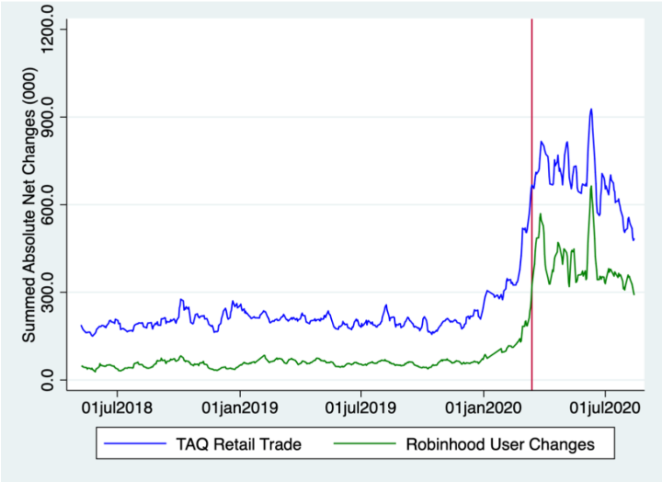


Fig. 2. Robinhood and TAQ Retail Trading

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*

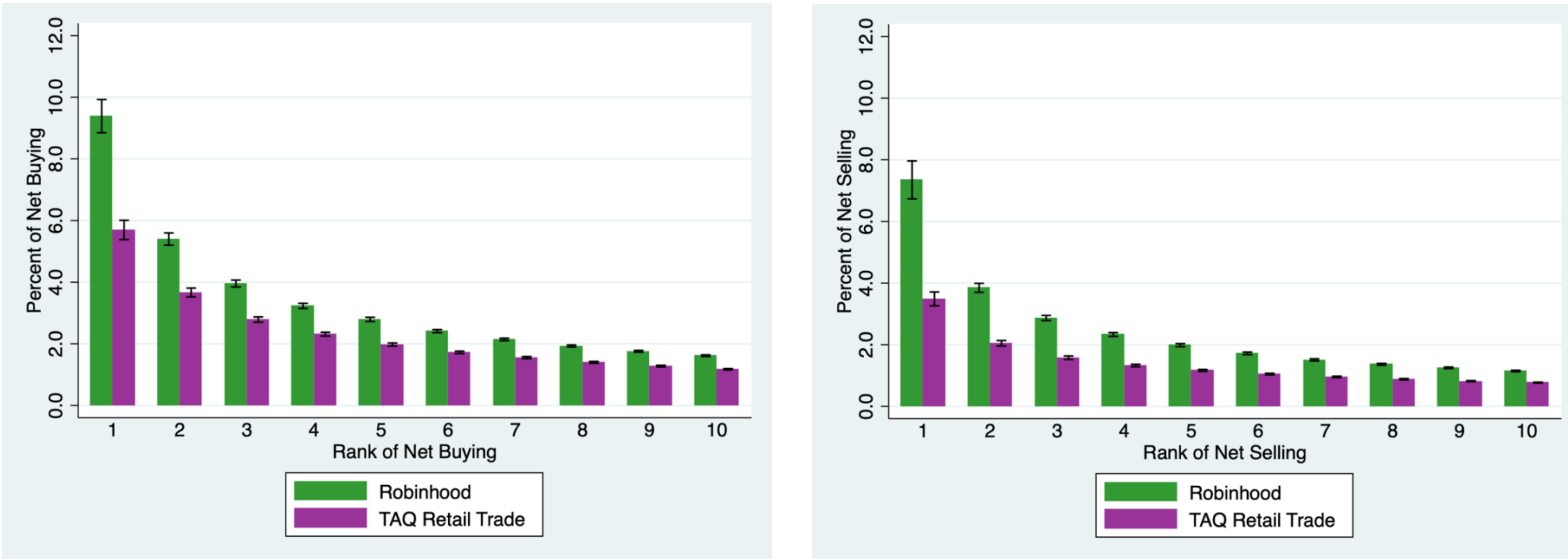
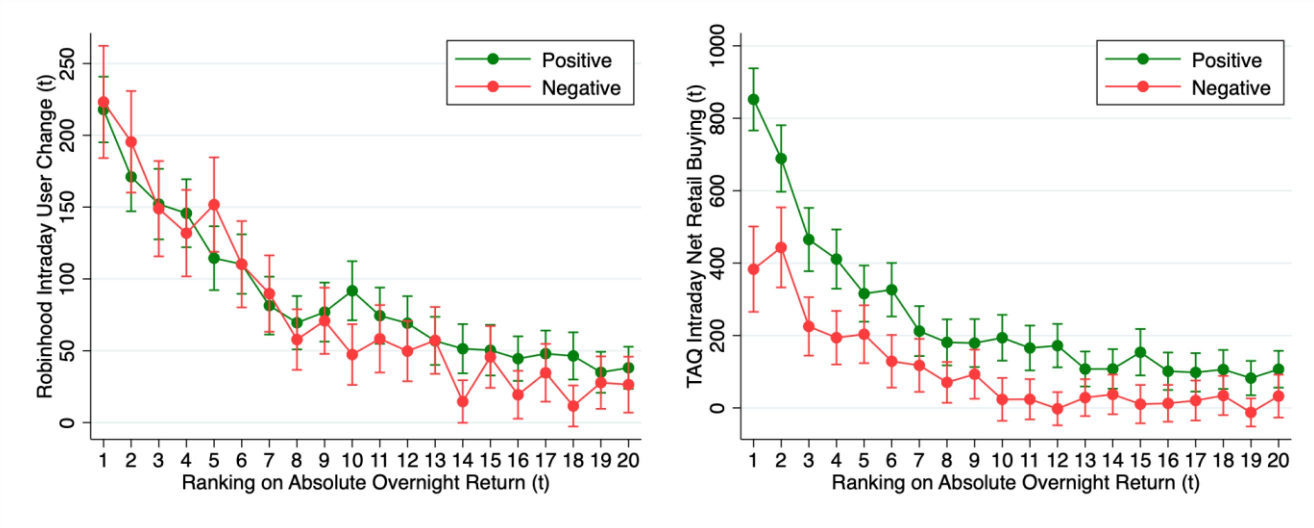


Fig. 3. The Concentration of Buying and Selling

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: Top Mover Rankings Sorted on Absolute Overnight Return



Panel B: Top Mover Rankings Sorted on Absolute Daily Return

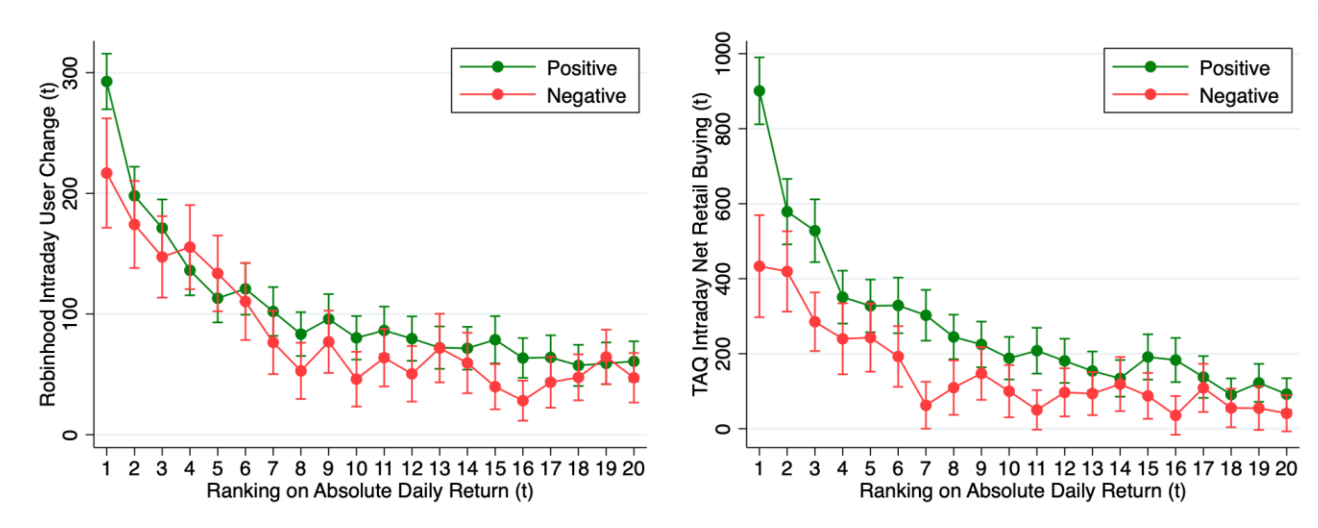
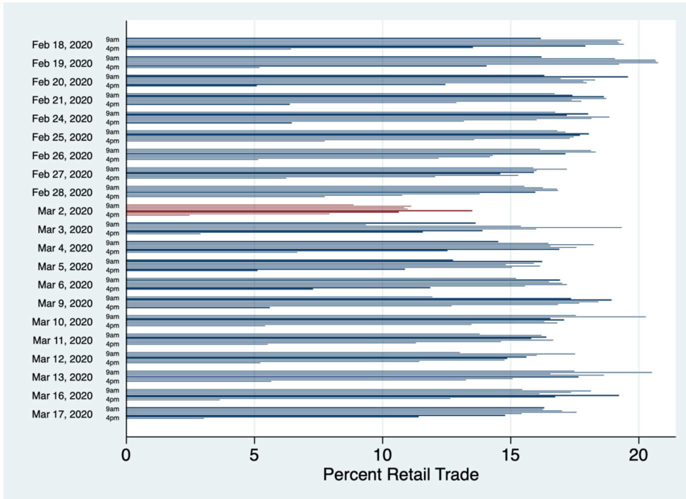


Fig. 4. Mean Robinhood (RH) Intraday User Change vs. TAQ Intraday Net Retail Buying on Top Mover Rankings

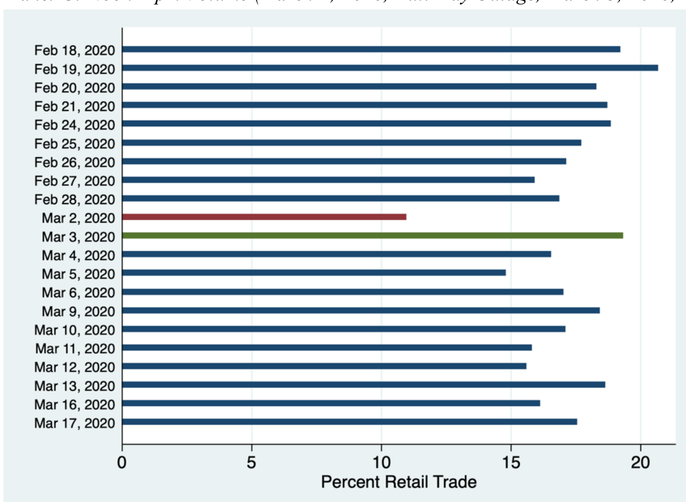
The figure presents the mean Robinhood intraday user change and TAQ intraday net retail buying on different measures top mover rankings. Panel A sorts top movers on absolute overnight return; Panel B sorts top movers on absolute daily return. RH Intraday User Change measures the change in RH users from first time stamp that excludes market open trades to the last time stamp before market closes, after adjusted for day fixed effects. TAQ Intraday Net Retail Buying is TAQ retail buying minus TAQ retail selling during the market trading hours, after adjusted for day fixed effects. The error bar represents the 95% confidence interval for the mean.

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)

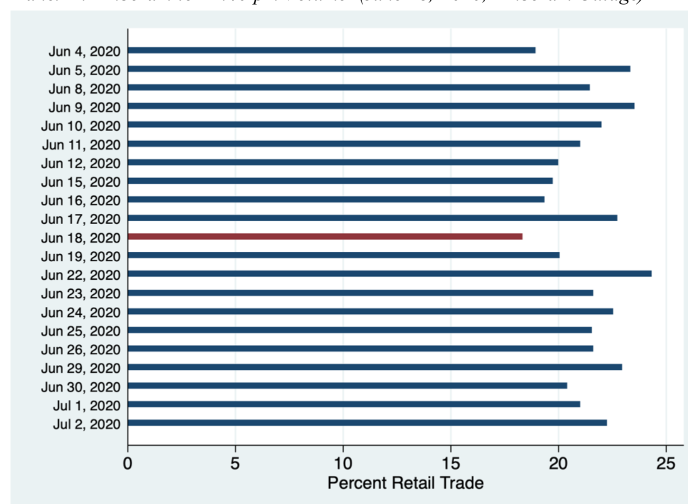


Fig. 5. The Effect of Robinhood Outages on Retail Trade

In panels A to C, the sample consists of the 50 most popular Robinhood stocks on Feb. 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 (red bar). A second shorter outage occurred around 10:00 am on March 3, 2020, and lasted for a bit more than an hour (red bar). In Panel C, March 2, 2020, is the day of a Robinhood outage (red bar). Robinhood tweeted all systems were fully restored at 11:55 am on March 3, 2020 (green 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 (red bars).

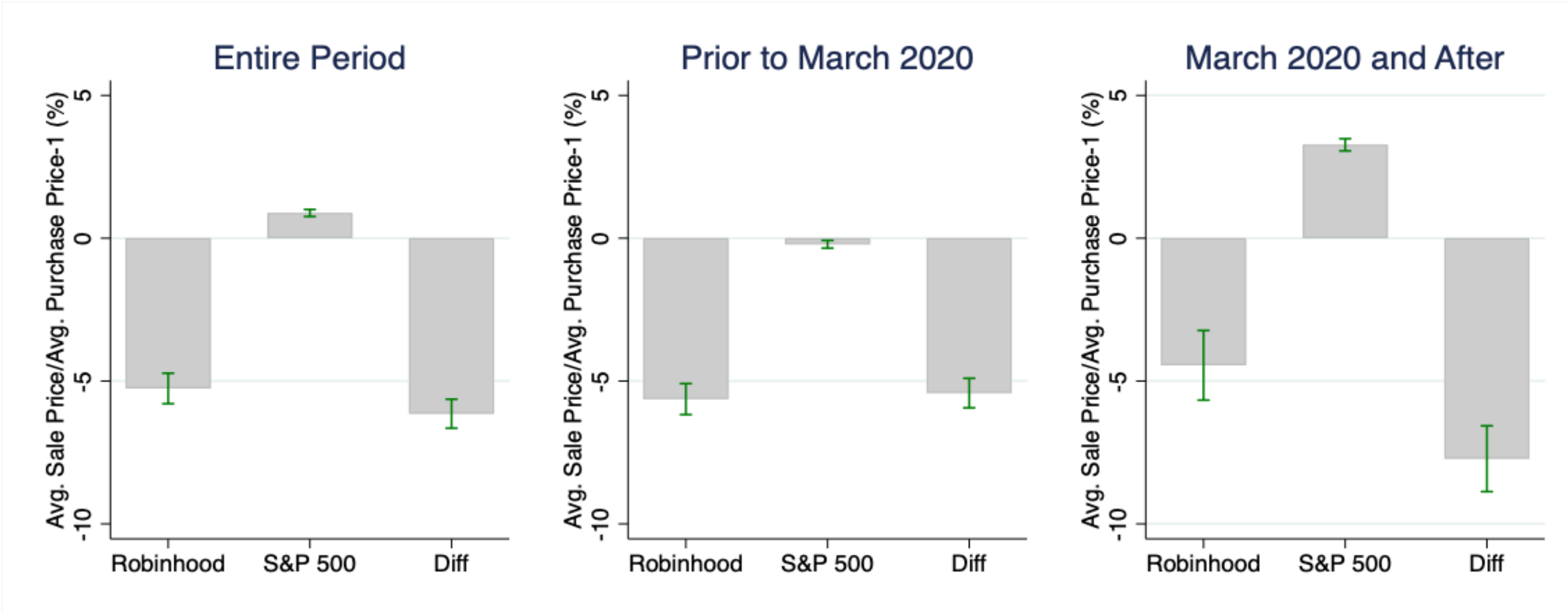


Fig. 6. Aggregate Investor Experience: Average Profitability

The figure presents the average profitability of Robinhood 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 calucated as (average sales price/average purchase price - 1). A positive profitability indicates the Robinhood community profited from that herding episode. The counterfactual trades on S&P 500 ETF assume 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.

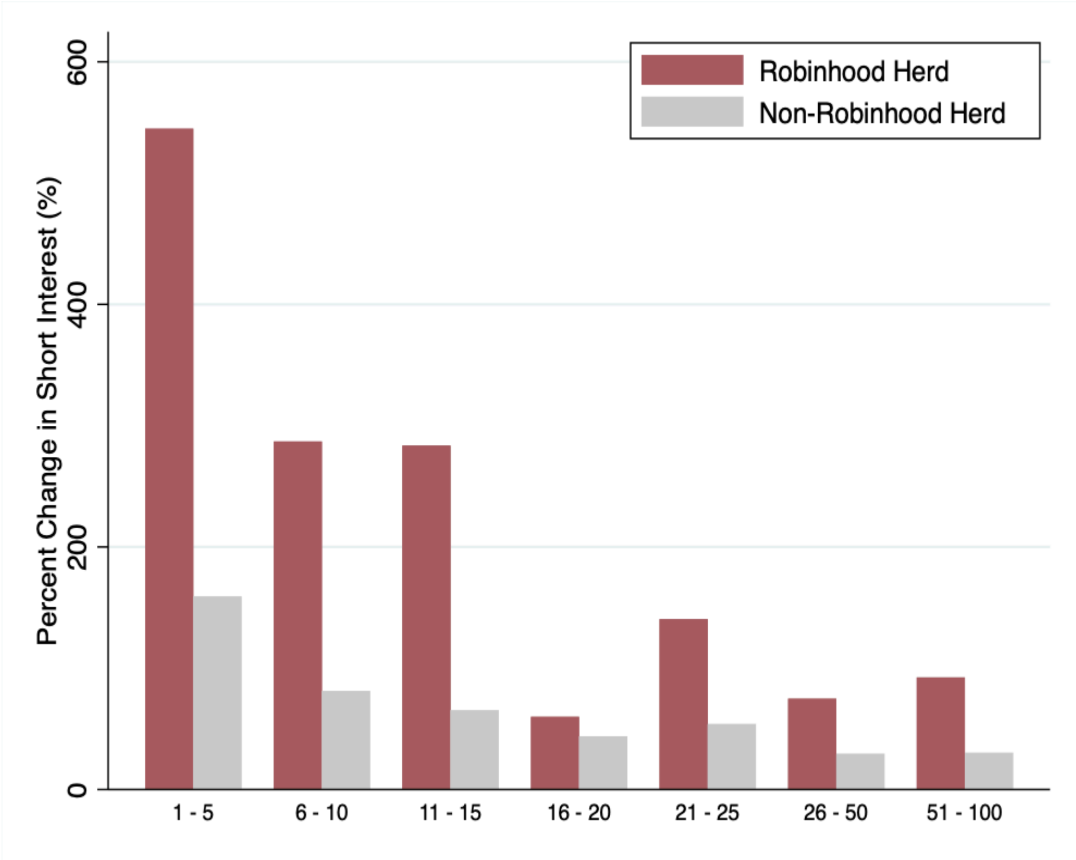


Fig. 7. Short Interest Changes and Robinhood Trading

The figure depicts the mean percent change in short interest. Short interest is measured bi-monthly using data from the NASDAQ. The x-axis is the maximum daily return ranking during the bi-monthly period. RH Herd are those stocks that are within the top 0.5% of user change percent at least one day within the bi-monthy period.

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

Panel A defines 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. Panel B defines herding events as securities with a user change ratio of at least 1.5 and a user change of at least 1000. 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), and *stocks* (number of stocks with users reported on Robintrack).

variable	N	mean	sd	min	p25	p50	p75	max
Panel A: Robinhood Securities in Top 0.5% of Positive User Change Ratio								
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 ($mil)	4,299	2,231.98	11,532.14	0.11	45.29	375.29	1,229.05	468,894.20
ret (%)	4,711	14.02	52.58	-91.79	-7.10	4.88	20.56	874.84
openret (%)	4,707	10.99	39.19	-88.60	-0.94	2.04	11.61	563.90
dayret (%)	4,710	3.43	30.68	-74.69	-7.79	-0.06	8.02	595.19
daily_buys	4,675	3,498.46	9,536.55	0.00	252.00	774.00	2,669.00	162,678.00
daily_sells	4,675	2,710.01	7,108.48	0.00	205.00	615.00	2,149.00	110,145.00
net_buys	4,675	788.44	2,962.92	-8,873.00	7.00	98.00	487.00	56,504.00
taq_retimb	4,675	0.11	0.17	-0.75	0.01	0.09	0.20	1.00
Panel B: Robinhood Securities with User Change Ratio > 1.5 and User Change ≥ 1000								
users_close	900	8,396.14	13,645.97	1,114.00	2,697.50	3,912.50	7,772.00	145,566.00
users_last	900	8,882.98	14,252.43	1,114.00	2,843.50	4,121.00	8,198.00	152,073.00
userchg	900	4,415.31	6,948.69	1,000.00	1,431.50	2,150.50	4,184.00	85,193.00
userratio	900	4.35	12.12	1.50	1.73	2.17	3.53	257.34
prc	866	15.38	38.99	0.22	2.60	6.12	14.61	828.34
size ($mil)	821	2,283.30	10,602.05	0.62	33.18	292.35	945.29	148,367.60
ret (%)	865	41.86	95.42	-91.79	-3.28	19.29	56.74	874.84
openret (%)	865	32.93	69.73	-88.60	0.75	10.10	44.09	563.90
dayret (%)	866	9.15	51.85	-74.69	-12.50	-0.83	15.71	595.19
daily_buys	858	12,182.55	18,438.14	0.00	2,196.00	5,179.50	13,477.00	162,678.00
daily_sells	858	9,060.49	13,870.25	0.00	1,486.00	3,615.50	10,631.00	110,145.00
net_buys	858	3,122.07	5,788.32	-8,873.00	443.00	1,305.00	3,135.00	56,504.00
taq_retimb	858	0.17	0.15	-0.43	0.07	0.14	0.25	0.66

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. 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 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 consists the top 10 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 during the market trading hours. *Top mover score* assigns a score for each rank, with 20 for the highest absoluate return and 1 for the 20th highest. *Negative return* is an indicator variable which 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	7.741***	30.16***	8.292***	28.63***
(=20: highest absret; =1: 20th highest absret)	(0.416)	(1.535)	(0.425)	(1.480)
Negative return	23.33**	-228.5***	-12.77	-190.0***
(=1: top mover return is negative; =0, otherwise)	(9.948)	(27.403)	(10.318)	(28.363)
Top mover score X Negative return	1.847***	-11.39***	-0.967	-11.59***
	(0.701)	(2.016)	(0.715)	(2.088)
Day FE	Yes	Yes	Yes	Yes
Observations	10084	10084	10176	10176
R-squared	0.251	0.247	0.235	0.246

Table 5. Determinants of Robinhood Herding Indicator Variable (Top 0.5% Percentage User Change)

The table exmines the determinants that predict the Robinhood herding indicator variable (top 0.5% percentage user change) using a linear probability model. rh_herd (t-1) is the lagged Robinhood herding indicator. Extreme absolute return (t-1) is an indicator which equals one if the absolute return is ranked in the top 10 on day t-1. Absolute Return (t-1) is the absolute value of the stock return on day t-1. Negative Return (t-1) is an indicator which equals to one if the return on day t-1 is negative. User Change (t-1) is the change in Robinhood users from day t-2 to day t-1. ln(Users(t-1)) is the logrithm of Robinhood uses before market closes on day t-1. Abnormal Vol (t-1) is the logrithm of the ratio of stock market volume on day t-1 to the average volume from day t-50 to t-11. Robust standard errors are clustered on day and stock level. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)
Dep var:		*rh_herd*	
rh_herd (t-1)	0.103***	0.107***	0.107***
	(0.005)	(0.005)	(0.005)
Extreme absolute return (t-1)	0.0666***		
(=1: top 10 absret; =0: otherwise)	(0.004)		
Absolute Return (t-1)		0.0673***	0.0664***
		(0.005)	(0.007)
Negative Return (t-1)			-0.000424**
(=1: ret<0; =0: otherwise)			(0.000)
Absolute Return (t-1)			0.00311
X Negative Return			(0.011)
User Change (t-1) (in 000s)	0.00845***	0.00708***	0.00707***
	(0.001)	(0.001)	(0.001)
ln (Users (t-1))	0.000183***	0.0000531***	0.0000536***
	(0.000)	(0.000)	(0.000)
Abnormal Vol (t-1)	0.000535***	0.000284***	0.000284***
= ln(Vol(t-1)/AvgVol(t-50, t-11))	(0.000)	(0.000)	(0.000)
Observations	3792584	3792584	3792584
R-squared	0.022	0.021	0.021

Table 6. 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 times 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 1, Table XX). 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 the all systems back online sometime between 11 am and noon; observations are stock-hours. 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 spans June 4 to July 2. Outage is an indicator variable that takes a value of one for the time intervals between 11:35 and 12:35, 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

	(1)	(2)	(3)
	All Stocks	50 Popular Stocks	50 High Attention Stocks
Panel A: March 2 Outage (all day)			
Outage	-0.723***	-5.227***	-4.366***
	(0.203)	(0.338)	(0.306)
Observations	1,090,382	8,395	8,363
R-squared	0.393	0.745	0.669
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***	-6.476***
	(0.172)	(0.712)	(0.796)
Repair	1.581***	3.742***	6.944***
	(0.133)	(0.131)	(1.236)
Observations	1,038,326	7,997	5,216
R-squared	0.397	0.764	0.188
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***	-1.393***
	(0.067)	(0.455)	(0.261)
Partial	0.476***	1.005***	0.033
	(0.091)	(0.312)	(0.197)
Repair	0.731***	1.543***	0.385***
	(0.092)	(0.190)	(0.043)
Observations	9,443,439	81,814	71,622
R-squared	0.231	0.621	0.212
Day FE	YES	YES	YES
Ticker FE	YES	YES	YES
Time of Day FE	YES	YES	YES

Table 7. 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. Abornal 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 comupted by clustering on event day. *% Positive* is the percent of returns that are positive. Panel A reports results for the top 0.5% of buys (*rh_herd*) whereas Panel B reports results for the extreme herding events (*rh_herd2*
*** p<0.01, ** p<0.05, * p<0.1

Event Day	AR	Std. Err.	% Positive	BHAR	Std. Err.	% Positive
Panel A: Herding Events (Top 0.5% Percentage User Change)						
			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%

Event Day	AR		Std. Err.	% Positive	BHAR		Std. Err.	% Positive
Panel B: Extreme Herding Events (At Least 50% Increase in Users and 1000 New Users)								
Pre-Event								
-10	0.32%		0.44%	44%	0.32%		0.44%	44%
-9	0.09%		0.37%	47%	0.54%		0.72%	44%
-8	0.13%		0.37%	47%	0.58%		0.69%	46%
-7	-0.32%		0.38%	43%	0.18%		0.77%	46%
-6	-0.38%		0.33%	43%	-0.26%		0.88%	45%
-5	-0.20%		0.42%	46%	-0.54%		1.00%	46%
-4	0.90%		0.54%	49%	0.34%		1.17%	48%
-3	1.27%		0.80%	47%	2.55%		2.23%	50%
-2	0.43%		0.48%	48%	2.43%		1.69%	51%
-1	15.71%	***	2.01%	63%	21.05%	***	3.48%	59%
0	42.03%	***	3.31%	73%	73.43%	***	7.01%	71%
Post-Event								
1	-2.17%	***	0.80%	38%	-2.17%	***	0.80%	38%
2	-1.43%	***	0.53%	40%	-3.93%	***	0.87%	35%
3	-0.82%		0.52%	42%	-5.12%	***	0.84%	33%
4	-0.47%		0.43%	42%	-5.91%	***	0.82%	33%
5	-0.28%		0.44%	40%	-6.39%	***	0.85%	33%
6	-0.89%	**	0.38%	40%	-7.00%	***	1.00%	32%
7	-0.20%		0.43%	43%	-7.38%	***	1.08%	33%
8	0.57%		0.54%	44%	-7.06%	***	1.01%	33%
9	0.42%		0.42%	44%	-6.87%	***	1.10%	32%
10	0.57%		0.46%	42%	-6.58%	***	1.10%	33%
11	-0.28%		0.41%	41%	-7.27%	***	1.00%	33%
12	0.07%		0.35%	43%	-7.35%	***	1.07%	33%
13	-0.12%		0.35%	44%	-7.59%	***	1.11%	32%
14	-0.32%		0.31%	43%	-7.98%	***	1.10%	31%
15	-0.21%		0.28%	44%	-8.36%	***	1.08%	30%
16	0.34%		0.30%	47%	-8.18%	***	1.13%	32%
17	-0.15%		0.33%	42%	-8.16%	***	1.28%	32%
18	-0.04%		0.29%	42%	-7.89%	***	1.53%	32%
19	-0.33%		0.31%	42%	-8.39%	***	1.48%	32%
20	0.34%		0.54%	44%	-8.45%	***	1.42%	32%

Table 8. Calendar Portfolio Returns of Herding Events

The dependent variable is the dollar weighted average daily returns over the risk-free rate. For each *rh_herd* 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. 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 independent variable is *alpha*, which is in basis points. 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)	(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	-0.601***	-0.557***	-0.540***	-0.519***	-0.509***	-0.494***	-0.944***	-0.826***	-0.785***
	(0.112)	(0.109)	(0.107)	(0.117)	(0.118)	(0.117)	(0.305)	(0.285)	(0.273)
Mkt_Rf	0.812***	0.711***	0.662***	0.654***	0.519***	0.446***	0.881***	0.726***	0.699***
	(0.110)	(0.120)	(0.127)	(0.120)	(0.130)	(0.135)	(0.138)	(0.155)	(0.156)
SMB		0.695***	0.453**		0.486	0.353		0.715**	0.393
		(0.217)	(0.219)		(0.308)	(0.303)		(0.317)	(0.384)
HML		0.092	0.345		-0.426*	-0.156		0.290	0.528
		(0.180)	(0.216)		(0.231)	(0.269)		(0.311)	(0.394)
MOM		-0.119	-0.227		-0.499***	-0.584**		0.045	-0.088
		(0.158)	(0.165)		(0.233)	(0.237)		(0.232)	(0.266)
RMW			-0.954***			-0.967***			-0.948
			(0.301)			(0.337)			(0.595)
CMA			-0.775*			-0.692			-0.873
			(0.424)			(0.521)			(0.755)
Observations	558	558	558	451	451	451	107	107	107
R-squared	0.19	0.234	0.254	0.057	0.081	0.099	0.407	0.479	0.494

Table 9. Regression of Daily Returns on Lagged Robinhood Herding Indicator (Top 0.5% Percentage User Change)

The dependent variable is the close to open return (*openret* in columns 1-3), open to close return (*dayret* in columns 4-6), or close to close return (*ret* in columns 7-9), which are all winsorized at the 0.1% level. The key independent variable is *rh_herd*, which is an indicator variable that takes a value of one on day *t* if the percentage change in users is in the top 0.5% for the day among stocks with positive user changes and a minimum of 100 users on day *t-1*. Control variables include retail order imbalance from taq (*taq_retimb*), lagged returns (*ret*), (unreported) 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:	*openret*	*openret*	*openret*	*dayret*	*dayret*	*dayret*	*ret*	*ret*	*ret*
L.rh_herd	-0.182	-0.190*	-0.095	-1.229***	-1.221***	-1.080***	-1.336***	-1.344***	-1.110***
	(0.111)	(0.113)	(0.121)	(0.133)	(0.135)	(0.144)	(0.160)	(0.163)	(0.174)
L2.rh_herd	-0.129*	-0.134*	-0.167*	-0.653***	-0.646***	-0.651***	-0.758***	-0.763***	-0.798***
	(0.075)	(0.075)	(0.091)	(0.111)	(0.112)	(0.115)	(0.129)	(0.130)	(0.141)
L3.rh_herd	0.060	0.044	0.087	-0.363***	-0.336***	-0.277**	-0.298**	-0.287**	-0.183
	(0.069)	(0.070)	(0.076)	(0.107)	(0.106)	(0.112)	(0.122)	(0.122)	(0.132)
L4.rh_herd	0.120*	0.112*	0.160**	-0.445***	-0.432***	-0.435***	-0.318***	-0.314***	-0.266**
	(0.066)	(0.067)	(0.073)	(0.093)	(0.093)	(0.099)	(0.115)	(0.116)	(0.128)
L5.rh_herd	0.112**	0.117**	0.062	-0.402***	-0.396***	-0.386***	-0.231**	-0.226*	-0.273**
	(0.056)	(0.056)	(0.060)	(0.105)	(0.106)	(0.111)	(0.117)	(0.118)	(0.131)
L.taq_retimb		0.059***	0.060***		-0.017**	-0.016**		0.041***	0.043***
		(0.005)	(0.006)		(0.007)	(0.007)		(0.008)	(0.008)
L2.taq_retimb		0.029***	0.028***		-0.015**	-0.015**		0.012	0.012
		(0.004)	(0.004)		(0.007)	(0.007)		(0.008)	(0.008)
L3.taq_retimb		0.018***	0.020***		-0.019***	-0.018**		-0.003	0.000
		(0.005)	(0.005)		(0.007)	(0.007)		(0.008)	(0.008)
L4.taq_retimb		0.026***	0.028***		-0.019***	-0.018***		0.006	0.009
		(0.004)	(0.005)		(0.007)	(0.006)		(0.008)	(0.008)
L5.taq_retimb		0.019***	0.019***		-0.011	-0.010		0.007	0.009
		(0.004)	(0.005)		(0.007)	(0.007)		(0.008)	(0.008)
L.ret			-0.016*			-0.021**			-0.037***
			(0.008)			(0.009)			(0.012)
L2.ret			0.001			-0.003			-0.002
			(0.007)			(0.007)			(0.012)
L3.ret			-0.006			-0.014*			-0.021*
			(0.007)			(0.007)			(0.012)
L4.ret			-0.013*			-0.003			-0.017*
			(0.007)			(0.006)			(0.010)
L5.ret			0.003			-0.009			-0.005
			(0.006)			(0.006)			(0.010)
Observations	3,590,567	3,306,712	3,306,222	3,591,636	3,306,716	3,306,225	3,656,926	3,313,045	3,312,553
R-squared	0.247	0.261	0.262	0.098	0.104	0.105	0.196	0.203	0.205
Days	550	550	550	550	550	550	550	550	550
5-day AR	-0.019	-0.051	0.047	-3.092***	-3.032***	-2.829***	-2.942***	-2.934***	-2.630***
	(0.192)	(0.194)	(0.212)	(0.296)	(0.295)	(0.300)	(0.322)	(0.323)	(0.348)

Table 10. Regression of Daily Returns on Lagged Robinhood Herding Indicator Variable (at least 50% increase in users and 1000 new users)

The dependent variable is the open returns (openret in columns 1-3), open to close return (dayret in columns 4-6), or close to close return (ret in columns 7-9), which are all winsorized at the 0.1% level. The key independent variable is *rh_herd2*, which is an indicator variable that takes a value of one on day *t* if the change in users from day *t* to *t-1* is greater or equal to 1000 and the ratio of users on day *t* to users on day *t-1* greater of equal to 1.5. Control variables include retail order imbalance from taq (*taq_retimb*), lagged returns (*ret*), (unreported) lags of an indicator variable if the *rh_herd2* measure is missing (and rh_herd2 is set equal to zero), and day fixed effects. 5-day AR is the sum of the coefficients on the five lags of *rh_herd2*. 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:	*openret*	*openret*	*openret*	*dayret*	*dayret*	*dayret*	*ret*	*ret*	*ret*
L.rh_herd2	0.156	0.133	0.313	-3.003***	-2.968***	-2.768***	-2.316***	-2.370***	-1.995***
	(0.375)	(0.383)	(0.468)	(0.426)	(0.430)	(0.475)	(0.490)	(0.502)	(0.604)
L2.rh_herd2	-0.492*	-0.500*	-0.930**	-1.024***	-0.975***	-1.239***	-1.527***	-1.514***	-2.210***
	(0.296)	(0.288)	(0.452)	(0.345)	(0.341)	(0.354)	(0.432)	(0.422)	(0.557)
L3.rh_herd2	-0.053	-0.103	0.015	-0.419	-0.360	-0.113	-0.516	-0.513	-0.145
	(0.288)	(0.282)	(0.347)	(0.353)	(0.358)	(0.408)	(0.448)	(0.443)	(0.583)
L4.rh_herd2	-0.083	-0.107	0.130	-0.549*	-0.484*	-0.543	-0.617*	-0.583*	-0.395
	(0.223)	(0.226)	(0.322)	(0.281)	(0.283)	(0.345)	(0.347)	(0.347)	(0.527)
L5.rh_herd2	0.354	0.372	0.065	-0.845***	-0.809**	-0.782**	-0.436	-0.381	-0.668
	(0.261)	(0.261)	(0.349)	(0.322)	(0.324)	(0.374)	(0.422)	(0.422)	(0.554)
L.taq_retimb		0.073***	0.077***		-0.010	-0.008		0.063**	0.069***
		(0.018)	(0.016)		(0.016)	(0.013)		(0.027)	(0.024)
L2.taq_retimb		0.059***	0.053***		-0.006	-0.008		0.052*	0.043*
		(0.021)	(0.017)		(0.015)	(0.012)		(0.031)	(0.025)
L3.taq_retimb		0.023*	0.025**		-0.015	-0.012		0.005	0.011
		(0.012)	(0.011)		(0.015)	(0.013)		(0.022)	(0.018)
L4.taq_retimb		0.033**	0.035***		-0.015	-0.015		0.018	0.019
		(0.016)	(0.013)		(0.015)	(0.013)		(0.027)	(0.022)
L5.taq_retimb		0.022*	0.018		-0.029*	-0.027*		-0.007	-0.008
		(0.013)	(0.012)		(0.018)	(0.015)		(0.024)	(0.022)
L.winret			-0.026			-0.025			-0.051*
			(0.020)			(0.015)			(0.028)
L2.winret			0.027			0.024			0.051
			(0.022)			(0.016)			(0.033)
L3.winret			0.005			-0.021			-0.015
			(0.020)			(0.017)			(0.031)
L4.winret			-0.025			0.006			-0.020
			(0.019)			(0.015)			(0.029)
L5.winret			0.021			-0.007			0.015
			(0.017)			(0.016)			(0.027)
Observations	3,590,567	3,306,712	3,306,222	3,591,636	3,306,716	3,306,225	3,656,926	3,313,045	3,312,553
R-squared	0.002	0.003	0.011	0.002	0.002	0.005	0.004	0.004	0.010
Days	550	550	550	550	550	550	550	550	550
5-day AR	-0.117	-0.204	-0.407	-5.841***	-5.596***	-5.445***	-5.412***	-5.362***	-5.413***
	(0.811)	(0.814)	(1.126)	(0.868)	(0.876)	(0.982)	(1.108)	(1.109)	(1.580)

Table 11. Robustness of Post-Herding Return Patterns

The table presents the five-day abnormal return from models 7 to 9 of Table X for various subsamples or robustness checks. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)
Dep var:	*ret*	*ret*	*ret*
Panel A: Full Sample			
5-day AR	-2.942***	-2.934***	-2.630***
Std. Err.	(0.322)	(0.323)	(0.348)
Panel B: Only Herding Events with Negative Same-Day Return			
5-day AR	-2.428***	-2.432***	-2.459***
Std. Err.	(0.486)	(0.488)	(0.480)
Panel C: Quote Midpoint Returns			
5-day AR	-2.797***	-2.790***	-2.408***
Std. Err.	(0.322)	(0.324)	(0.354)
Panel D: Only stock with prices > $5			
5-day AR	-0.851***	-0.833***	-0.749**
Std. Err.	(0.242)	(0.244)	(0.265)
Panel E: Small Cap (< $1 billion in market cap)			
5-day AR	-4.250***	-4.254***	-3.835***
Std. Err.	(0.418)	(0.421)	(0.432)
Panel F: Large Cap (> $1 billion in market cap)			
5-day AR	0.14	0.135	-0.0508
Std. Err.	(0.534)	(0.538)	(0.521)
Panel G: Other than Common Stocks			
5-day AR	-3.174***	-3.149***	-2.940***
Std. Err.	(0.583)	(0.586)	(0.667)
Panel H: Post-Covid (after March 13, 2020)			
5-day AR	-4.581***	-4.691***	-3.073***
Std. Err.	(0.680)	(0.674)	(0.743)
Panel I: Pre-Covid (before March 13, 2020)			
5-day AR	-2.221***	-2.179***	-2.231***
Std. Err.	(0.337)	(0.341)	(0.330)

Table 12. Regression of Short Interest Changes on Returns and Top Robinhood Changes

The dependent variable is the change in short interest from t to t-1, where intervals are every two weeks. The top 0.1% of observations are winsorized. The key independent variables are rh_chgratio, which is the maximum one day percentage change in users during the period from t-1 to t, and rh_herd, which is an indicator variable that takes a value of one if the percentage change in users is in the top 0.5% during any day during the period from t-1 to t. Control variables include the maximum daily return (Max_ret) over the period t-1 to t and dummy variables representing the maximum daily return rank (Max_rank) over the t-1 to t period. 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:	Change in Short Interest			
rh_chgratio	157.32***		209.24***	
	(34.62)		(43.82)	
rh_herd		101.58***		125.51***
		(26.06)		(31.30)
Max_ret	350.17***	316.09***		
	(61.52)	(57.97)		
Max_rank_1			682.62***	662.08***
			(139.01)	(139.34)
Max_rank_2			259.79**	259.78**
			(108.88)	(109.20)
Max_rank_3			96.82*	79.78*
			(48.96)	(46.22)
Max_rank_4			115.77**	108.94**
			(53.33)	(53.42)
Max_rank_5			101.52**	83.67
			(51.48)	(55.99)
Max_rank_6			215.86	197.5
			(174.42)	(168.16)
Max_rank_7			12.97	5.27
			(14.42)	(17.20)
Max_rank_8			53.98*	51.37
			(31.85)	(31.21)
Max_rank_9			53.60	51.07
			(48.08)	(47.45)
Max_rank_10			43.38**	37.66*
			(20.23)	(21.28)
Max_rank_11-25			26.58*	20.74
			(15.00)	(14.64)
Max_rank_26-50			-4.21	-7.42
			(5.46)	(5.21)
Max_rank_51_100			-1.31	-2.60
			(5.69)	(5.48)
Observations	59,458	59,458	59,458	59,458
Avg. R-squared	0.034	0.032	0.039	0.036

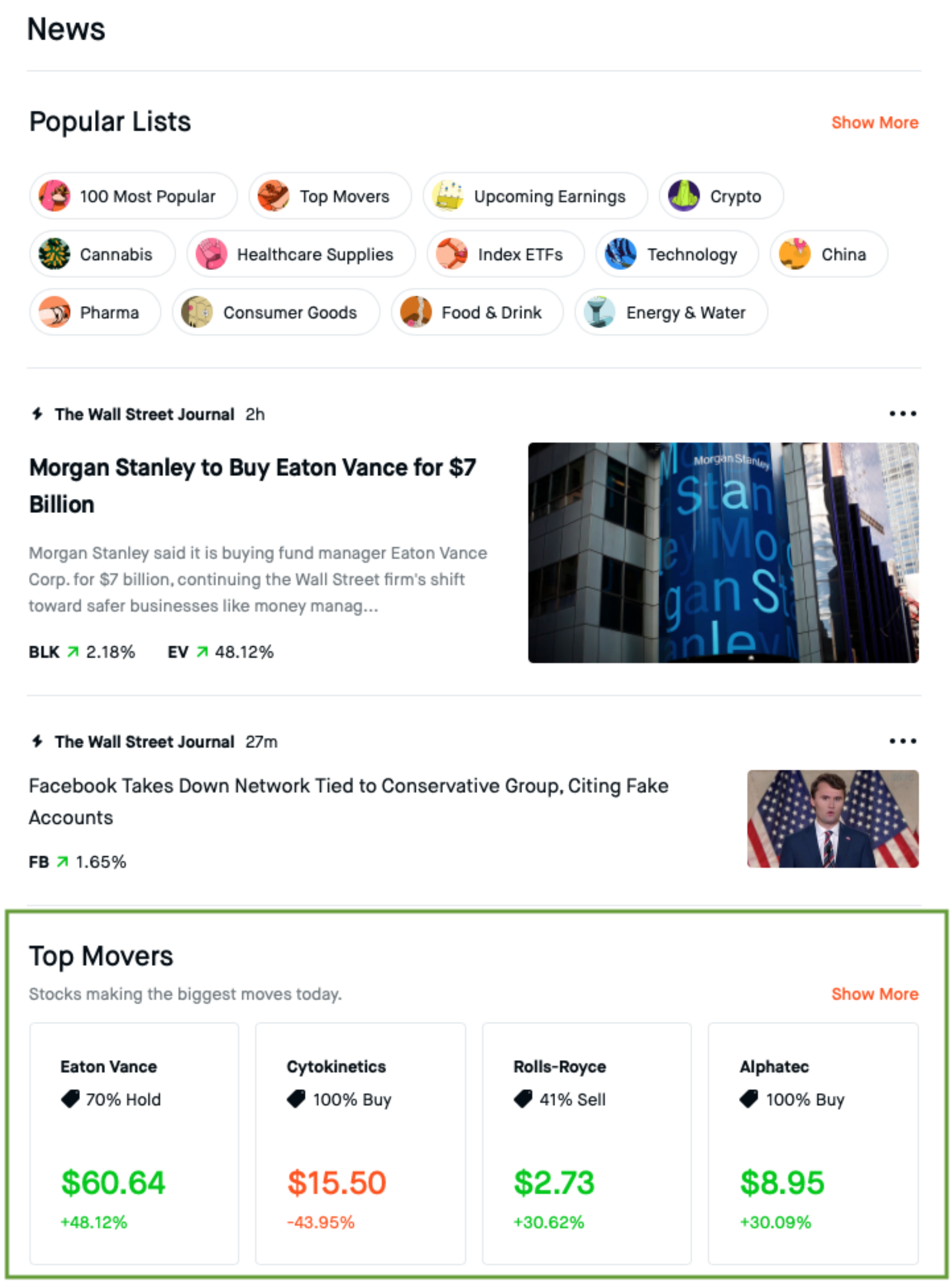


Fig. A1. Screenshot of Robinhood User Interface

The figure presents a screenshot of the interface of the "News" section from Robinhood website on Oct 8th, 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., Oct 7th, 2020 for 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.

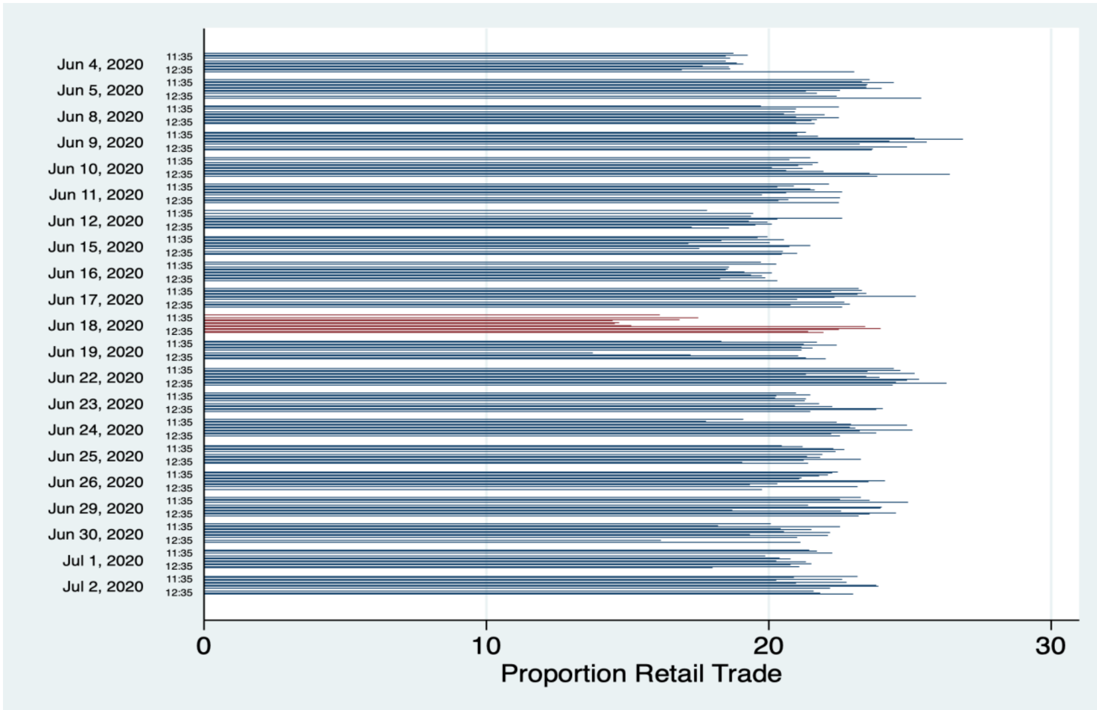


Fig. A2. Percent Retail Trading in Five-Minute Intervals during June Outage

This figure depicts retail volume at five-minute intervals for the same period depicted in Figure 8, Panel D (11:35am to 12:40pm).

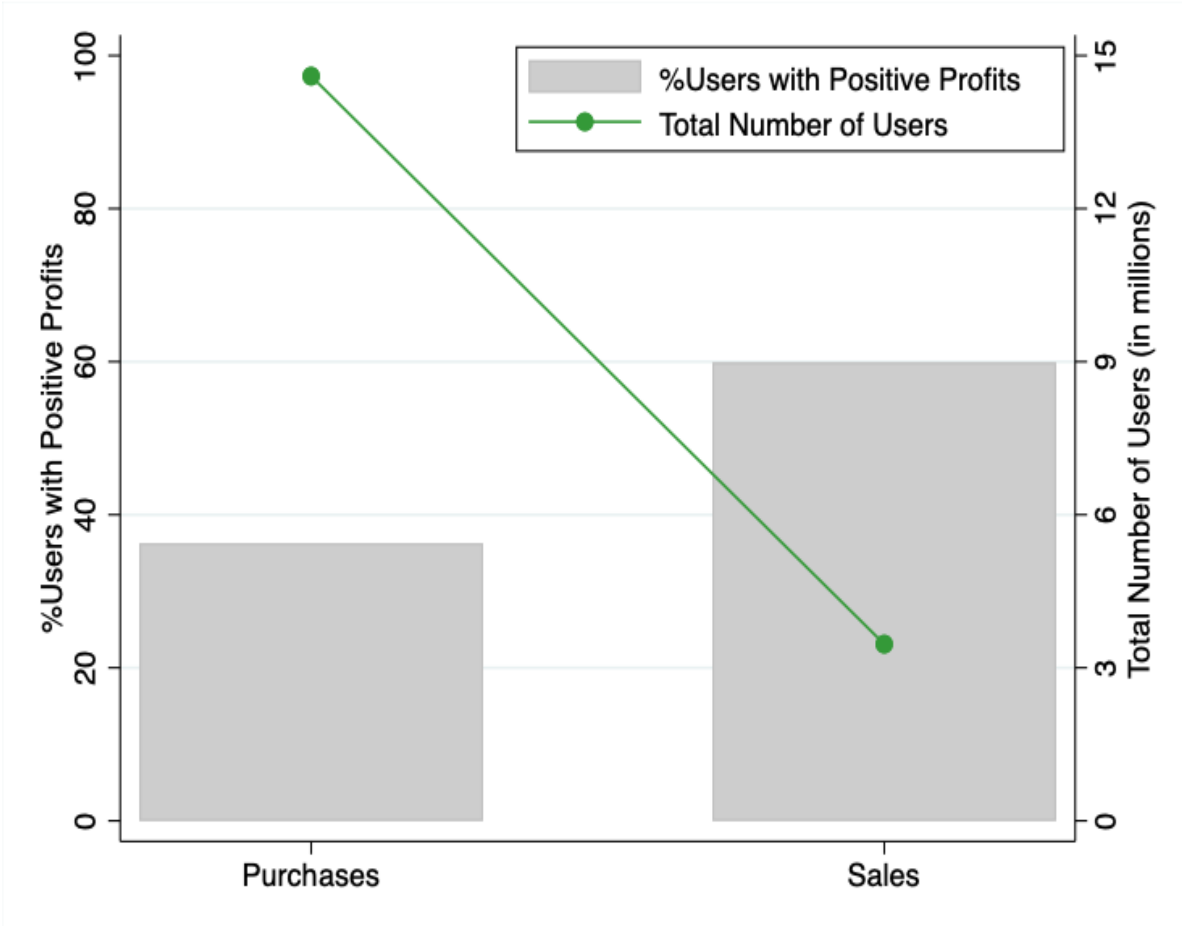


Fig. A3. Aggregate Investor Experience: Percentage of Investors with Positive Profits

The figure presents the average percentage of investors that 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 riskfree rate) on a portfolio that mimics the Robinhood user experience. We assume each new user buys $1/P shares of stock, update portfolio weights daily, and 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 independent variable is *alpha*, which is in basis points. 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)	(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)
Mkt_Rf	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***
		(0.233)	(0.244)						
RMW			-0.253**			-0.222***			-0.438*
			(0.106)			(0.072)			(0.232)
CMA			-0.605***			-0.300***			-0.545**
Observation	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. 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), and stocks (number of stocks with users reported on Robintrack).

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 A3. 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 (rh_negherd) is an indicator variable that takes a value of one on day t if the percentage change in users is in the top (bottom) 0.5% for the day among stocks with positive (negative) user changes 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*	*rh_negherd*
L.rh_herd	0.116***	0.192***
	(0.005)	(0.007)
L2.rh_herd	0.003	0.131***
	(0.002)	(0.006)
L3.rh_herd	0.010***	0.025***
	(0.002)	(0.004)
L4.rh_herd	0.008***	0.019***
	(0.002)	(0.003)
L5.rh_herd	0.002	0.009***
	(0.001)	(0.003)
L.rh_negherd	0.004***	0.128***
	(0.002)	(0.006)
L2.rh_negherd	0.004**	0.045***
	(0.002)	(0.004)
L3.rh_negherd	0.006***	0.030***
	(0.002)	(0.004)
L4.rh_negherd	0.006***	0.016***
	(0.002)	(0.003)
L5.rh_negherd	0.006***	0.022***
	(0.002)	(0.003)
Observations	3,949,293	3,949,293
R-squared	0.014	0.097

Table A4. Regression of Daily Returns on Lagged Robinhood Buy/Sell Herding Indicator (Top/Bottom 0.5% Percentage User Change)

The dependent variable is the open returns (*openret* in columns 1-3), open to close return (*dayret* in columns 4-6), or close to close return (*ret* in columns 7-9), which are all winsorized at the 0.1% level. The key independent variable is *rh_herd (rh_negherd)*, which is an indicator variable that takes a value of one on day *t* if the percentage change in users is in the top (bottom) 0.5% for the day among stocks with positive (negative) user changes and a minimum of 100 users on day *t-1*. Control variables include retail order imbalance from taq (*taq_retimb*), lagged returns (*ret*), (unreported) 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 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:	*openret*	*openret*	*openret*	*dayret*	*dayret*	*dayret*	*ret*	*ret*	*ret*
L.rh_negherd	-0.321***	-0.304***	-0.318***	0.067	0.053	0.084	-0.227**	-0.219*	-0.200
	(0.065)	(0.065)	(0.071)	(0.098)	(0.099)	(0.103)	(0.112)	(0.115)	(0.125)
L2.rh_negherd	-0.001	-0.007	0.055	-0.109	-0.098	-0.053	-0.116	-0.107	0.002
	(0.061)	(0.061)	(0.071)	(0.089)	(0.091)	(0.093)	(0.106)	(0.108)	(0.116)
L3.rh_negherd	-0.114**	-0.104*	-0.090	0.006	-0.007	-0.011	-0.081	-0.086	-0.075
	(0.055)	(0.055)	(0.057)	(0.081)	(0.081)	(0.086)	(0.100)	(0.101)	(0.106)
L4.rh_negherd	-0.050	-0.050	-0.103*	-0.034	-0.039	-0.020	-0.087	-0.090	-0.127
	(0.050)	(0.050)	(0.054)	(0.095)	(0.096)	(0.097)	(0.101)	(0.102)	(0.102)
L5.rh_negherd	-0.021	-0.018	-0.017	-0.116	-0.110	-0.151*	-0.121	-0.110	-0.150
	(0.049)	(0.050)	(0.055)	(0.080)	(0.081)	(0.085)	(0.084)	(0.086)	(0.094)
L.rh_herd	-0.190*	-0.198*	-0.103	-1.226***	-1.217***	-1.076***	-1.340***	-1.348***	-1.112***
	(0.111)	(0.112)	(0.121)	(0.133)	(0.135)	(0.144)	(0.160)	(0.164)	(0.175)
L2.rh_herd	-0.066	-0.075	-0.104	-0.669***	-0.659***	-0.668***	-0.717***	-0.724***	-0.759***
	(0.076)	(0.076)	(0.088)	(0.114)	(0.114)	(0.116)	(0.130)	(0.130)	(0.139)
L3.rh_herd	0.107	0.089	0.123	-0.352***	-0.325***	-0.279**	-0.242*	-0.234*	-0.153
	(0.072)	(0.073)	(0.077)	(0.108)	(0.107)	(0.111)	(0.125)	(0.125)	(0.130)
L4.rh_herd	0.156**	0.147**	0.182**	-0.434***	-0.420***	-0.429***	-0.275**	-0.272**	-0.244*
	(0.068)	(0.069)	(0.074)	(0.095)	(0.095)	(0.100)	(0.117)	(0.118)	(0.130)
L5.rh_herd	0.153***	0.156***	0.106*	-0.392***	-0.384***	-0.380***	-0.185	-0.179	-0.228*
	(0.059)	(0.059)	(0.062)	(0.106)	(0.107)	(0.112)	(0.122)	(0.122)	(0.133)
Observations	3,590,567	3,306,712	3,306,222	3,591,636	3,306,716	3,306,225	3,656,926	3,313,045	3,312,553
R-squared	0.247	0.261	0.262	0.098	0.104	0.105	0.196	0.203	0.205
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	-0.507	-0.483	-0.473	-0.186	-0.201	-0.151	-0.632	-0.612	-0.550
	0.121	0.122	0.118	0.177	0.179	0.181	0.213	0.219	0.215
5-day Buy Herd AR	0.160	0.119	0.205	-3.072	-3.005	-2.832	-2.759	-2.757	-2.496
	0.199	0.200	0.217	0.302	0.302	0.306	0.330	0.331	0.351