Fast and Furious: An Intraday Analysis of Robinhood Users' Trading Behavior*

David Ardia[a], Clément Aymard[a,b,*], Tolga Cenesizoglu[b]

[a]GERAD & Department of Decision Sciences, HEC Montréal, Montréal, Canada
[b]Department of Finance, HEC Montréal, Montréal, Canada

Preliminary

Abstract

We analyze the reaction of retail investors to price changes at the intraday level using hourly data on the number of investors holding a given stock on the Robinhood platform. We observe that (i) Robinhood users tend to open new positions on stocks exhibiting extreme returns in the past hours, (ii) this response to extreme returns is asymmetric, in the sense that it is more pronounced for stocks that experienced sharp declines in the last several hours, and, (iii) the speed of reaction is relatively fast with the net number of Robinhood investors opening a position in a given stock within an hour following a large negative return of that stock. Using proxy variables representing these patterns of Robinhood investors' reaction to past returns, we identify their determinants in the time-series and cross-section. In the time series, we find that market volatility or investor sentiment can influence the day-to-day behavior of RH investors. In the cross-section, firm characteristics such as size, level of debt, or systematic risks are important determinants. Further analyses reveal that these results may vary before and after the market correction associated to the COVID-19 pandemic announcement.

Keywords: Behavioral Finance, COVID-19, FinTech, Investor Attention, Intraday, Retail investors, Robinhood
JEL: D9, G11, G4

*We are grateful to IVADO and the Swiss National Science Foundation (grant #179281) for their financial support.

*Corresponding author. HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montreal, QC H3T 2A7. Phone: +1 514 340 6103.

Email addresses: `david.ardia@hec.ca` (David Ardia), `clement.aymard@hec.ca` (Clément Aymard), `tolga.cenesizoglu@hec.ca` (Tolga Cenesizoglu)

Preprint submitted to SSRN *February 17, 2022*

1. Introduction

Retail participation in stock markets is soaring. Between 2019 and today, the number of Americans having an account with one of the seven largest brokers has skyrocketed, from 59 to 95 million. In terms of trading flows, the part coming from retail investors peaked at more than 40% during the first quarter of 2021, from only 25% in mid-2018 (see, *e.g.*, The Economist, 2021). The online trading platform offering zero trading commissions, "Robinhood Market Inc.", is a prominent player driving this trend. Robinhood (hereafter RH) has been a spectacular success. Since its launch in 2013, the number of users has continued to grow, reaching 22.4 million in the third quarter of 2021, more than three times its level in 2018 (see, *e.g.*, BusinessofApps, 2021; Statista, 2021).

Not only does the number of users increase, but so does their importance in the markets. There are many examples of Robinhood users moving the market. For instance, on May 22, 2020, Hertz's rental car company filed for bankruptcy. On May 26, its stock price fell sharply, to $0.50. A wave of RH users bought the stock in the following days, triggering a massive rally. The stock price reached $5.53 on June 8, a return of more than 900% in less than two weeks. Another example took place in early 2021, with the now-famous Gamestop stock: Through discussions in Reddit's chat *r/wallstreetbets*, RH traders have come together to buy shares in an attempt to punish some big and influencing institutional investors who had shorted the stock. The enormous buying wave pushed Gamestop's price to crazy highs, forcing the institutional actors to close their short positions and take their losses.

With its mission to "democratize finance for all", Robinhood has attracted a new type of investors. They are typically young, tech-savvy, and inexperienced. They hold on average a tiny portfolio, with an estimated average account size between $3,000 and $5,000, and they trade mainly through the RH's smartphone app, which has often been criticized for its design prone to the "gamification" of investing.[1]

Hence, it is not surprising to find a growing literature analyzing the behavior of this new type of investor. For example, Barber et al. (2021) and Welch (2021) explore some aspects of the behavior of Robinhood investors at the daily frequency. However, none of these studies analyze the *intraday* behavior of these investors. In this paper, we fill this gap and analyze the reaction of RH investors to price movements at an intraday scale. This finer scale allows us to answer different questions about the speed of their reaction to extreme price movements, which is not possible using low-frequency data. This intraday analysis is crucial in understanding these investors who grew up in the internet and social media era and are ultra-connected. Besides, as advocated by Barber et al. (2021), they are heavy users of the RH's app, which can send notifications anytime. Therefore, it seems very likely that they adjust their positions more than once a day.

[1] This typical RH investor profile has been drawn multiple times in press articles (see, *e.g.*, New York Times, 2020; CNN Business, 2020; CNBC Markets, 2020; Forbes Advisor, 2021; CNBC Markets, 2021; Barron's, 2021), or mentioned in scientific papers (see, *e.g.*, Barber et al., 2021; Welch, 2021; Eaton et al., 2021; Van der Beck and Jaunin, 2021; Chapkovski et al., 2021).

In a first set of analyses, using a comprehensive dataset from Robintrack.net and high-frequency prices from the NYSE Trade And Quote millisecond datasets (MTAQ), we explore how RH investors react to past returns.[2] At the aggregate level (*i.e.*, all stocks, whole period), we observe three interesting facts.

First, RH investors tend to react positively (*i.e.*, they tend to open new positions) and strongly to stocks exhibiting extreme returns in the past hours. In other words, if we group past returns by level (from lowest to highest returns), we observe that RH investors' reaction to these returns is U-shaped. This result is consistent with some findings from the existing literature on retail investors. For example, Barber and Odean (2008), using private data from brokerage firms, also observe this U-shape. In contrast to this study, our results focus only on RH investors. Most importantly, our results are observed intra-daily. Welch (2021), using extreme daily returns, also finds that "RH investors in 2020 liked to purchase both large gainers and large losers". Here again, our results differ because we look at RH investors' behavior on a finer scale.

Second, we observe that this response to past extreme returns is asymmetric. RH investors tend to open more positions on stocks exhibiting large negative returns than on stocks exhibiting large positive returns.[3] RH investors appear to be more contrarian-style investors or "buy-the-dip" investors. This observation could also be related to some behavioral biases identified in the literature such as the disposition effect (Shefrin and Statman, 1985), which is the tendency to sell past winners and hold on past losers.

The third key observation relates to the speed of reaction of RH investors to past extreme returns. We find that RH investors respond more strongly during the first hour for stocks exhibiting large past negative returns and increasingly weakly in the following hours. In other words, the intensity of the reaction decreases monotonically as the hours pass. However, this is not true for the RH investors' reaction to past returns from different group levels (*i.e.*, moderate returns or large positive returns). For these returns, the strength of RH investors' reaction is not significantly different after one or more (up to five) hours. A possible interpretation of this result is that (i) RH investors' attention is mainly directed towards stocks exhibiting past sharp declines, and (ii) they tend to act on this information reasonably fast.

The second set of analyses is devoted to exploring the variations of these three RH investors' behavioral patterns in the time-series and cross-section.

In time-series analyses, we first evaluate how RH investors react to past returns *on a daily basis* and observe significant variations. In an attempt to evaluate the determinants of these

[2]While other studies are devoted to finding the best proxy for the attention of individual investors, this is not the focus of this paper. We instead assume that investors with limited experience, like RH investors, are likely to be influenced by the most straightforward and visible information they are subject to: price movements. Papers studying other proxies of attention include, for instance, Kaniel and Parham (2017) who use news from the Wall Street Journal, Da et al. (2011) and Andrei and Hasler (2015) who advocate for Google search data, and Cookson and Niessner (2020) who exploit messages from the StockTwits platform.

[3]In their daily frequency work, Barber and Odean (2008) find mixed results: Asymmetry is more or less pronounced depending on the proxy (brokerage firms) used.

variations, we construct (daily) time-series of various proxy variables representing the RH investors' behavioral patterns described above (*i.e.*, reaction to extreme returns, asymmetry of reaction to extreme returns, speed of reaction), and regress them on time-series factors. The factors we test for are market volatility, market excess returns, and investors' sentiment. We find that market volatility is negatively correlated to RH investors' reaction to previous extreme returns and RH investors' speed of reaction to previous extreme negative returns. Therefore, RH investors tend to respond more weakly (more slowly) to stocks exhibiting extreme (extreme negative) returns on low-volatility days. Market excess returns appear to be a less important determinant in the variation of RH investors' reaction to extreme returns and RH investors' speed of reaction to extreme negative returns. However, it seems to influence the asymmetry of reaction of RH investors to extreme returns. On days of rising market, RH investors tend to respond more strongly to extreme negative returns but less so to extreme positive returns. The third factor, investor sentiment, seems to be a critical determinant of the reaction of RH investors to extreme returns, particularly to extreme negative returns. When the sentiment is positive, RH investors react more strongly. It also correlates positively with the speed of RH investors' reaction to extreme negative returns. We also perform analyses to evaluate how the announcement of the COVID-19 as a global pandemic affect our results. Interestingly, we observe that market volatility – and to a lesser extent, investors sentiment – influences RH investors' behavior differently in the pre- and post-COVID announcement period. Before the announcement, market volatility had a more substantial effect on RH investors' reaction to extreme returns and RH investors' speed of reaction to large negative returns. We also find that, in the pre-COVID period, investors sentiment was positively and strongly correlated to both RH investors' reaction to extreme returns and RH investors' speed of reaction to large negative returns. However, this effect disappears after the COVID-19 market correction of March 2020.

Similarly, we evaluate how these patterns compare at the company level. As for the time-series analyses, we compute our proxy variables representing the key RH investors' behavioral patterns *on a company basis* and observe significant cross-sectional variations. To identify the determinants, we regress these proxies on various firm characteristics, namely firm size, book-to-market ratio, level of debt, trading volume, liquidity, systematic risk, and idiosyncratic risk, while controlling for firm sectors. Among other results, we find that firm's size affects RH investors' reaction to previous extreme returns differently depending on the sign of the returns. When larger firms exhibit extreme negative returns, RH investors tend to react more strongly by opening more positions in these firms. However, when larger firms exhibit extreme positive returns, RH investors tend to react more weakly as they open fewer positions in these firms. Therefore, the asymmetry of reaction to extreme returns (toward large negative returns) is more pronounced for firms with large market capitalization. Regarding the reaction speed to extreme negative returns, we find that RH investors tend to react more quickly to larger firms exhibiting past large negative returns. The firm's level of debt also appears to be an important determinant. For firms with more leverage, RH investors tend to react more strongly when they post extreme returns, their reaction is more asymmetric toward large negative return, and their speed of reaction is higher. Whether systematic or idiosyncratic, the risk level is also positively correlated to our proxies of RH investors' behavior. The riskier the company, the stronger the reaction of RH investors to previous extreme returns, the greater the asymmetry in their reaction to extreme returns in the direction of negative returns, and the faster their reaction to extreme negative returns.

Finally, we investigate the effect of the COVID-19 pandemic announcement on these results. We find that for some "popular" industries (*e.g.* Consumer Discretionary, Information Technology, Health Care), the reaction asymmetry is less pronounced after the announcement. Also, the speed of reaction to extreme negative returns tends to slow down after the announcement for these sectors. In addition, we observe that splitting the effects of the firm characteristics and industries before and after the COVID-19 pandemic announcement improves the explanatory power considerably for most of the key RH investors' behavioral patterns. This suggests that the market correction following the COVID-19 pandemic announcement date might have significantly impacted the trading behavior of RH investors. Indeed, some firm characteristics have a different impact on RH investors' behavior in the pre- and post-COVID sub-period. For example, we observe that after March 11, 2020, RH investors were quicker to respond to big-capitalization firms exhibiting past large negative returns, but it was not the case before this date. Another interesting finding relates to the effect of the firm's systematic risk. Before the announcement, RH investors responded more quickly to higher-betas firms exhibiting past extreme negative returns, but after the announcement, the relation is no longer significant.

The rest of the paper is organized as follows. Section 2 presents the related literature and our contributions. Section 3 discusses our data sources and defines our main variables. Section 4 presents the three key findings related to RH investors' reaction to past returns. Sections 5 and 6 examine the determinants of the behavior of these RH investors in time series and cross-section, respectively, and analyze how these results differ before and after the announcement of the COVID-19 pandemic. Section 7 concludes.

2. Related literature

Our research contributes to the broad literature studying retail investors. Black (1986) is among the first to take an interest in these investors and found that they mostly "trade on noise", a conclusion somewhat confirmed by more recent studies of Kumar and Lee (2006) or Fong et al. (2014). The identification of these noise traders gave rise to an extensive literature on retail investor sentiment (see, *e.g.*, Baker and Wurgler, 2006; Stambaugh et al., 2012), which we relate to through our time-series analyses. In contrast to a major strand of this literature that focuses on individual investors' performance (see, *e.g.*, Barber et al., 2009a; Coval et al., 2021; Gargano and Rossi, 2018), or their asset pricing implications (see, *e.g.*, Barber et al., 2009b; Kelley and Tetlock, 2013; Kaniel et al., 2008; Hvidkjaer, 2008), we do not treat these topics. Instead, we rather concentrate on individual investors' attention and trading behavior. In particular, given the profile of RH investors described above, we choose a simple attention proxy —previous extreme returns at the hourly frequency— and examine the trading behavior of RH investors conditional on this proxy. While more scarce, notable studies treating this more specific topic include Barber and Odean (2008), who use news from DJ News Service, unusual volume and extreme returns to proxy for attention and investigate its impact on investors' trading using data from brokerage accounts or Yuan (2015), who uses extreme returns of the DJIA and news from the New York Times and Los Angeles Times to proxy for attention. To proxy for investors' trading, he uses all orders below a certain size. Our paper is different because we investigate the impact of extreme returns on a brand new type of retail investor, the RH investors. In addition, we evaluate the reaction through an intraday perspective, which has never been done before to the best of our knowledge.

We also contribute to the burgeoning literature studying RH investors specifically. These studies use the same data source as us, from Robintrack.net. Welch (2021) constructs a representative RH portfolio (the "ARH" portfolio) and analyzes its composition and performance at the daily frequency. Interestingly, he finds that RH investors tend to hold stocks with above-average trading volume and purchase both large gainers and large losers and that this portfolio did not underperform with respect to traditional models (*i.e.*, risk-free rate, market model, Fama-French five-factor plus momentum models). This paper mainly focuses on performance. While it gives some insights about RH investors' behavior following extreme returns, it does not provide such a comprehensive study as we do. Most importantly, it uses data at the daily frequency. Another major paper from Barber et al. (2021) shows that herding events (periods of intense buying by RH investors) imply subsequent negative returns. They also find that during RH outages (period where RH investors can not trade due to technical breakdowns), the overall activity of retail investors[4] is significantly reduced. This shows evidence that RH investors represent a large chunk of total individual investor activity, reinforcing the relevance of a more in-depth study of their behavior like the one we provide here. They also highlight that the RH app's unique features impact the way they trade. Other papers on RH investors, less closely linked to our study but with important results, are worth noting. Eaton et al. (2021) examine the effect of RH users' trading activity on liquidity. They found that during RH outages, market liquidity of RH-favored stocks improves, and their volatility decreases. Friedman and Zeng (2021) also make use of RH outages and find that during such events, retail activity is reduced (consistent Barber et al. (2021) findings) and bid-ask spreads narrow (consistent with Eaton et al. (2021) findings). They also show that RH investors' activity is higher for stocks whose prices are more responsive to earnings surprises. Moss et al. (2020) shows that RH investors did not particularly care about ESG investing. Ozik et al. (2021) show that during the COVID-19 pandemic

[4] To identify retail buys and sells, they used the Boehmer et al. (2021) algorithm.

lockdown in Spring 2020, RH investors activity sharply increased, consistent with the fact that the new fintech platforms are favoring retail investors participation as they can trade from home. In this paper, we also assess the impact of COVID-19 pandemic locking on RH investors but through a different lens: We examine if the lockdown affected the *strenght* of reaction to extreme returns through time-series analyses. Ben-David et al. (2021) find that "sentiment-driven investors" like RH investors are particularly prone to invest in thematic ETFs. Van der Beck and Jaunin (2021) developed a structural model to quantify the impact of Robinhood traders on the US equity market.

3. Main data and variable definitions

3.1. Robinhood holdings and intraday prices

From mid–2018 to mid–2020, it was possible to use Robinhood's API to obtain information on the number of RH investors holding a particular stock at a given time. In agreement with Robinhood, the website Robintrack.net (hereafter, RT) developed a script to continuously pull down the information, which was then shared online on RT's website. In August 2020, Robinhood restricted access to its API (see, *e.g.*, Fortune, 2020). This paper uses RT's data files that contain observations representing the number of RH investors holding security i with an associated timestamp $t_{i,k}$, with k representing the kth observation on a total of K_i observations for stock i. We denote this variable $N_{i,t_{i,k}}$. The original timestamp associated to each observation does not correspond to the time of the observation, but the time at which the data has been extracted from Robinhood.com. Because the extraction process implies a 45-minute delay, we substract 45 minutes from each original timestamp to obtain the real timestamp.[5] A clear definition of the nature of the data is important. An increase in the variable $N_{i,t_{i,k}}$ indicates that there were more users who opened a position in that stock compared to those who completely closed their positions. However, the change in the number of users does not provide any information on how much is held in each account. It might well be the case that the number of shares of a given stock held by all RH users decreases even if the change in the number of users is positive. However, we believe that the change in the number of users provide some useful information on the demand for this stock by RH users.[6] Also, as indicated by Welch (2021), this measure can be influenced by RH's referral program which offers free shares to investors opening a new account on the platform.

[5]Our discussion with Casey Primovic, the administrator of Robintrack, led us to estimate the timestamp lag between 30 and 60 minutes. Therefore, a data point with an associated time of 10:45 actually represents a snapshot of the actual data point 30 to 60 minutes before that time. In this paper, we assume a timestamp lag of 45 minutes. For robustness check, all results have been replicated using timestamp lags of 30 and 60 minutes (see appendix). A discussion of this caveat is also provided in Barber et al. (2021).

[6]For example, if the number of RH users for Apple Inc. (AAPL) increases on a given time interval, it means that the number of RH users who *opened* a new position in AAPL is greater than the number of RH users who *closed* an existing position in AAPL. It does not mean that the total number of shares of AAPL held by RH investors is higher, as it is possible that some RH investors reduced or added to an already existing position in the stock during that time interval.

Our second main data source is the NYSE Trade And Quote millisecond datasets (MTAQ) from which we obtain intraday prices. For each stock, we match RT's observations ($N_{i,t_{i,k}}$) to the last price available of that stock before its real timestamp ($t_{i,k}$). For each stock, we also match RT's observations to the last price available of the SPDR S&P 500 ETF (SPY) – our market proxy – before its real timestamp ($t_{i,k}$). For a consistent extraction of intraday prices through the MTAQ databases, we apply the filters described in the algorithm of Barndorff-Nielsen et al. (2009).[7]

The original RT's database spans the period from May 1, 2018 to August 13, 2020, and contains more than 140 million intraday observations, on more than 8,000 distinct securities. To ensure data quality, we apply some adjustments before and after the match with transaction prices. In particular, we follow Welch (2021) and drop the first month of the original period as we suspect the reliability of the data improves as time passes. To match RT's data and trade prices, we reduce the number of observations to market-opening hours (9.30 a.m. to 4.00 p.m.) only. We also retain only common-stocks type securities (CRSP share codes of 10 or 11). For a detailed list of the adjustments and their impact on the original dataset, we refer to Section A.1 of the appendix.

Our final sample contains 8.03 million observations on 2,853 stocks and 527 trading days from June 1st, 2018 to August 13, 2020.

3.2. Variable definitions

Our main variable of interest is the change in the number of RH users holding a given stock between two consecutive observations. We use this variable to proxy for the aggregate trading behavior of RH users concerning a given stock. We define this variable as:

$$\Delta N_{i,t_{i,k}} = \log\left(\frac{N_{i,t_{i,k}}}{N_{i,t_{i,k-1}}}\right)\frac{60}{MNT(t_{i,k-1},t_{i,k})} \tag{1}$$

where $N_{i,t_{i,k}}$ refers to the RT's indicator defined earlier. Since two consecutive observations are not necessarily regularly spaced, we normalize the change in the number of RH users to an hourly level. Hence, the second term is a one-hour scaling factor where $MNT(t_{i,k-1},t_{i,k})$ is the number of minutes between the two consecutive stock i observations. Note that the change might occur between two intraday observations (when $t_{i,k-1}$ and $t_{i,k}$ are from the same day) or overnight (when $t_{i,k-1}$ and $t_{i,k}$ are from two different days). In the latter case, the variable expresses the change between the last observation available before the market closes (4.00 p.m.) on a given trading day $d-1$, and the first observation available after the market opens (9.30 a.m.) on the following trading day d.[8]

[7] In step "P3", we retain entries originating from the three main exchanges: NYSE, NASDAQ, and AMEX. We do not apply the filter "T4" because do not extract quote data.

[8] To avoid zeros in the denominator of the first term in the RHS of (1), we add one to all $N_{i,t_{i,k}}$ entries.

For a given stock, we define the return at timestamp $t_{i,k}$ as:

$$R_{i,t_{i,k}} = \log\left(\frac{p_{i,t_{i,k}}}{p_{i,t_{i,k-1}}}\right)\frac{60}{MNT(t_{i,k-1},t_{i,k})} \tag{2}$$

where $p_{i,t_{i,k}}$ is the price of stock i at time $t_{i,k}$. To better capture the true extreme movements, we scale the return by their volatility over the entire period. We define the adjusted returns as:

$$r_{i,t_{i,k}} = R_{i,t_{i,k}}/\hat{\sigma}(\{R_{i,t_{i,k}}\}), \tag{3}$$

where $\hat{\sigma}(\cdot)$ is the full-sample standard deviation estimator. When $t_{i,k}$ corresponds to the first timestamp of a day, the return is overnight and we adopt the approach by Lou et al. (2019) to compute it.[9] We also standardize the overnight returns by their full-sample standard deviation to obtain adjusted returns. These adjusted overnight returns are appended to the series of adjusted intraday returns.

Table 1 presents summary statistics on our RH trading behavior ($\Delta N_{i,t_{i,k}}$) and high-frequency adjusted return measure ($r_{i,t_{i,k}}$). The statistics are computed over all stock-day-time observations. The average (log) change in RH users is positive and rather high, at 4.36 basis points (BP). This is due to the success of Robinhood: the number of RH users was almost constantly increasing during our sample period. When a new user registers, he opens new positions to build his portfolio, which directly translates into a higher number of total users holdings. Because a large part of the observations do not change from one timestamp to another, the median change is zero. The distribution of this measure appears to be positively skewed, in particular due to extreme positive values.

Our high-frequency adjusted returns measure displays a mean of approximately 0 and a standard deviation of 1. While 75% of observations lie within ± 0.41, we note the presence of very large extreme returns, the minimum and maximum being -30.33 and $+44.42$ standard deviations away from the mean, respectively.[10]

4. The reaction of RH investors to high-frequency returns: empirical facts

In this section, we analyze the reaction of RH users to high-frequency adjusted returns. We start with a preliminary analysis using averages of RH users change conditional on stock return levels. We then tackle the question through regression analyses which allow to control for other factors potentially impacting RH investors' reaction.

[9]Specifically, we define $R_{i,t_{i,k}} = \log\left(\frac{p_{i,d\left(t_{i,\overline{k}}\right)}}{p_{i,d-1\left(t_{i,\overline{k}}\right)}}\right) - \log\left(\frac{p_{i,d\left(t_{i,\overline{k}}\right)}}{p_{i,d\left(t_{i,\underline{k}}\right)}}\right)$, where $d(t_{i,k})$ is the day associated to timestamp $t_{i,k}$, and $p_{i,d\left(t_{i,\underline{k}}\right)}$ and $p_{i,d\left(t_{i,\overline{k}}\right)}$ are respectively the first and last prices available on day d.

[10]More details on the RH users' change and high-frequency adjusted returns are provided in Section A.2 of the appendix.

4.1. Preliminary evidences

We start with a simple non-parametric approach to demonstrate the reaction of RH investors to stock returns of different signs and magnitude. We do this by grouping observations based on the previous returns. Specifically, we classify each volatility-adjusted returns $r_{i,t_{i,k}}$ into group levels. Formally, we define six groups of return levels that form a partition of $\mathbb{R}$: $\mathcal{G}_1 =]\infty, -2[, \mathcal{G}_2 = [-2, -1[$, $\mathcal{G}_3 = [-1, 0[, \mathcal{G}_4 = [0, 1[$, $\mathcal{G}_5 = [1, 2[$, $\mathcal{G}_6 = [2, \infty[$. Then we define the generic indicator function $I_{\mathcal{G}_g}(x)$ that is equal to one if $x \in \mathcal{G}_g$ and zero otherwise.

The measure we are interested in is the average RH users change at the h hour(s) horizon (*i.e.*, the number of hours elapsed between the observation of the return and the RH investors' reaction), conditional on the returns belonging to level group $\mathcal{G}_g$.[11] At this stage, we compute this measure on all stock-day-time observations. That is, we do not differentiate between stocks or sub periods. With horizons from one to five hours (h), and six groups of return levels (g), we estimate a total of 30 conditional averages.

Panel A, Figure 1 displays the results with respect to group of return levels in the horizontal axis for different horizons. At all horizons, the average change in RH users conditional on past returns is U-shaped: RH investors tend to open more new positions in stocks exhibiting large price movements in the previous hours, whether positive or negative, than in stock exhibiting moderate price movements in the previous hours. For example, the number of RH users increases on aggregate by 9.23BP following a large negative ($\mathcal{G}_1$) return in the previous hour, and the number of RH users increases on aggregate by 5.67BP following a large positive return ($\mathcal{G}_6$) in the previous hour. In comparison, the increase in the number of RH users following a moderate return (*e.g.*, $\mathcal{G}_4$) in the previous hour is only 0.33BP. This U-shape seems to be present across all horizons.

This panel also shows that RH investors' reaction to previous returns is asymmetric. Indeed, we actually observe a smirk: the average change in RH users following a large *negative* return is greater than the average change in RH users following a large *positive* return. Using the same example as above, we see that the average RH users change conditional on previous large negative return in the previous hour (9.23BP) is more than 60% superior to its counterpart conditional on past large positive return (5.67BP). The magnitude of the asymmetry, however, seems to decrease as the horizon increases. Indeed, at the five-hour horizon, the U-shape is not as asymmetric: the average RH users change conditional on a previous large negative return (6.09BP) is similar in magnitude to its counterpart conditional on a previous large positive return (5.54BP).

Panel C, Figure 1 displays the results with respect to the horizon in the horizontal axis for different group of return levels. This representation helps interpret results in terms of speed

[11] Formally, the conditional average can be expressed as:

$$\mu_{g,h} = \sum_i \sum_k \Delta N_{i,t_{i,k}} I_{\mathcal{G}_g}(r_{i,t_{i,k-h}}) \Big/ \sum_i \sum_k I_{\mathcal{G}_g}(r_{i,t_{i,k-h}}).$$

of reaction. The average reaction to large negative returns decreases monotonically as the horizon increases. It means that RH investors' reaction to extreme negative price movements is stronger in the first hour, and tends to fade as the hours pass. However, this phenomena is not observable for positive extreme returns. Interestingly, the conditional averages are also monotonically decreasing for the second lowest group of returns ($\mathcal{G}_2 = [-2, -1[$), but not for any other groups. Overall, these results suggest that RH investors tend to respond particularly fast to stock exhibiting past large negative returns by opening new positions in these stocks. However, it is difficult to draw conclusions about their speed of reaction relative to past positive returns, as the strength of the response is similar when assessed over the one to five hour horizon.[12]

4.2. Regressions analyses

One limitation of the results presented above is the lack of control for other factors that might impact the reaction of RH investors. In particular, when assessing the reaction of RH investors h hour(s) after a given return, one might want to control for the other returns of that stock occurring before or after the period of interest. We control for previous returns up to five hours. For example, if we are interested in the reaction of RH investors one hour after a given return ($h = 1$), we control for the returns two to five hours ($h = 2, \ldots, 5$) before this reaction. Similarly, if we assess the reaction of RH investors two hours after a given return ($h = 2$), we control for the returns one hour and three to five hours ($h = 1, 3, 4, 5$) before this reaction. Another potentially important factor is the market return as in some cases, RH investors' reaction could be market-wide and not stock specific. We therefore control for market returns as well, at one to five hours preceding the reaction of RH investors. We specify the regression as follows:

$$\begin{aligned}\Delta N_{i,t_{i,k-h}} = &\sum_{g=1}^{G} \beta_{g,h} r_{i,t_{i,k-h}} I_{\mathcal{G}_g} + \sum_{j \in J} \gamma_j r_{i,t_{i,k-j}} + \sum_{j \in J} \delta_j r^2_{i,t_{i,k-j}} \\ &+ \sum_{j \in J'} \psi_j r_{MKT,t_{i,k-j}} + \sum_{j \in J'} \xi_j r^2_{MKT,t_{i,k-j}} + \epsilon_{i,t_{i,k-h}} ,\end{aligned} \tag{4}$$

where $I_{\mathcal{G}_g}$ is the indicator function defined previously and G is the number of groups. $r_{i,t_{i,k-j}}$ are the other stock i returns preceding RH investors' reaction by up to five hours, where $J = \{x \in \{1, 2, \ldots, 5\} | x \neq h\}$. Similarly, $r_{MKT,t_{i,k-j}}$ represent the market-adjusted returns where $J' = \{1, 2, \ldots, 5\}$. To account for non-linearity, we also add quadratic versions of these controls, $r^2_{i,t_{i,k-j}}$ and $r^2_{MKT,t_{i,k-j}}$.

We are interested in the $\beta_{g,h}$ coefficients which represent the sensitivity of the change in RH users to stock returns belonging to a specific group of return level $\mathcal{G}_g$ at a given horizon h after controlling for the effect of other factors. We estimate five regressions, one for each horizon.

[12] Data associated to Figure 1 can be consulted in Section A.3 of the appendix.

Table 2 presents estimation results. All five regressions are based on the full panel (all stock-day-time observations) and are estimated by ordinary least squares. Incorporating the control variables do not significantly impact the results suggested by the approach using conditional averages. Out of the 30 estimates, 28 are significant at the 1% level. More important, all coefficients evaluating RH investors reaction to extreme returns ($\mathcal{G}_1$ and $\mathcal{G}_6$) are highly significant. The U-shape pattern is present at all horizons. For example, at the one-hour horizon, reactions to extreme returns controlled for other factors ($\beta_{1,1} = 7.95$ and $\beta_{6,1} = 2.71$) are stronger than reactions to moderate return ($\beta_{3,1} = 0.31$ or $\beta_{4,1} = -0.2$). The asymmetry pattern also manifests at the one-hour horizon as the average RH investor reactions to large negative price movements (7.95) is about 193% superior to the average reaction to large positive price movements (2.71). This asymmetry is however less pronounced as the horizon increases because at the five-hour horizon, the reactions to large positive and negative price movements are closer, at respectively $\beta_{1,5} = 3.57$ and $\beta_{6,5} = 2.78$). With regards to RH investors' speed of reaction to past extreme returns, our previous observations also hold: the coefficients with respect to the lowest group of return, $\mathcal{G}_1$, monotonically decreases as the horizon increases (from 7.95 at the one-hour horizon to 3.57 at the five-hour horizon). This monotonic decrease is also observable for the second strongest group of negative return ($\mathcal{G}_2$). For other groups of return levels, however, no patterns seem to emerge.

Panels B and D of Figure 1 illustrate these results. Although they are in line with our preliminary evidence, we note some differences. Within the regression setup, the U-shape is more asymmetric at all horizons, and the monotonic decrease in the average reaction to large negative returns as a function of horizon is more pronounced. Therefore, after controlling for other stock returns and market returns, RH investors' "pure" reaction to extreme negative returns is stronger than its counterpart with respect to extreme positive returns. Similarly, with these controls, RH investors' "pure" speed of reaction to large negative returns seems to be greater.

5. Time series variation in RH investors' behavior and its determinants

5.1. Daily reaction of RH investors to high-frequency returns

The results presented so far shed light on key patterns of RH investors' behavior with respect to previous returns at the *aggregate level* (*i.e.*, on all stock-day-time observations). Given the high-frequency nature of our sample, we can investigate time series variations by running regression (4) separately for each day. Our sample counts 526 trading days from June 1st, 2018 to August 13, 2020. Therefore, we run 526 daily regressions for each of the five horizons, resulting in a total of 2,630 estimations. We denote the coefficients of interest representing the average reaction of RH investors to a given group of return level g at a given horizon h and a given day d as $\beta_{g,h}^{(d)}$.

The average number of observations per regression is 15,239. Figure 2 displays these regressions' estimates corresponding to extreme groups of return levels (*i.e.*, $\beta_{1,h}^{(d)}$ and $\beta_{6,h}^{(d)}$).

We can observe that, for all horizons, the estimates exhibit important fluctuations over time. Therefore, the three RH investors' behavioral patterns described above – that RH investors respond stronger to extreme returns than to moderate returns, RH investors respond stronger to large negative returns than to large positive returns, and RH investors' speed of reaction is higher following large negative returns – must also vary over time.[13]

We now turn to the identification of factors driving these variations. We proceed in two steps. First, using these daily estimates, we construct proxy variables representing the three main RH investors' behavioral patterns identified in the previous section. Second, we regress these proxies on time-series factors.

5.2. Proxy variables

The first two patterns we observed at the aggregate level are that (i) RH investors tend to open new positions on stocks exhibiting large price movements in the past few hours and (ii) these responses are asymmetric, in the sense that they are of greater magnitude for stocks experiencing sharp declines in the last few hours. To measure these two patterns over time, we introduce the following variables:

$$Ext^{(d)} = \frac{\hat{\beta}^{(d)}_{g=1,h=1} + \hat{\beta}^{(d)}_{g=6,h=1}}{2} - \frac{\hat{\beta}^{(d)}_{g=3,h=1} + \hat{\beta}^{(d)}_{g=4,h=1}}{2} \tag{5}$$

$$ExtNeg^{(d)} = \hat{\beta}^{(d)}_{g=1,h=1} - \frac{\hat{\beta}^{(d)}_{g=3,h=1} + \hat{\beta}^{(d)}_{g=4,h=1}}{2} \tag{6}$$

$$ExtPos^{(d)} = \hat{\beta}^{(d)}_{g=6,h=1} - \frac{\hat{\beta}^{(d)}_{g=3,h=1} + \hat{\beta}^{(d)}_{g=4,h=1}}{2} \tag{7}$$

$$Asy^{(d)} = \hat{\beta}^{(d)}_{g=1,h=1} - \hat{\beta}^{(d)}_{g=6,h=1}\,. \tag{8}$$

While it is possible to compute these variables for each horizon, to keep results concise, we only focus on variations at the one-hour horizon (*i.e.*, $h = 1$)[14]. $Ext^{(d)}$ represents the average daily RH investors' reaction to all extreme returns ($\mathcal{G}_1$ and $\mathcal{G}_6$) relative to their average reactions to moderate returns ($\mathcal{G}_3$ and $\mathcal{G}_4$). $ExtNeg^{(d)}$ and $ExtPos^{(d)}$ split the reaction according to the sign of the extreme returns. The former evaluates RH investors' reaction to large negative returns only ($\mathcal{G}_1$) relative to the average reaction to moderate returns, and the latter evaluates RH investors' reaction to large positive returns only ($\mathcal{G}_6$) relative to the average reaction to moderate returns. The reason why we do this separation is to help explain the asymmetry of reaction to past extreme returns. Finally, the last proxy, $Asy^{(d)}$, is a more direct way of measuring this asymmetry. The higher the value of the indicator, the stronger the asymmetry of the reaction towards past large negative returns.

[13] For easier reading, only the coefficients of extreme group of return levels are displayed in Figure 2. For a time-series representation of the coefficients of all groups, and summary results on the estimation of these daily regressions, refer to section A.4 of the appendix.

[14] For a visualization of distributions of these proxies at horizons one to five, refer to section A.5 of the appendix

The third behavioral pattern we observed at the aggregate level relates to the speed of reaction of RH investors to past extreme negative returns. We measure it with the following proxy:

$$SpeedExtNeg^{(d)} = \hat{\beta}^{(d)}_{g=1,h=1} - \hat{\beta}^{(d)}_{g=1,h=5} \tag{9}$$

This indicator compares the strength of the reaction to previous extreme negative returns at the one-hour horizon, to the strength of the reaction to previous extreme negative returns at the five-hour horizon. The higher the value of the indicator, the faster RH investors respond (by opening more positions in the stock during the first hour than during the fifth hour) to previous negative returns.

Figure 4 displays the daily series of the five variables defined above. Not surprisingly, all variables exhibit fluctuations over time. For example, Panel A shows that on certain days, the average RH users change following extreme returns from the past hour is almost zero (*e.g.*, in October, 2018), but for other days it can peak to a value close to 15 (*e.g.*, in early 2020). Similarly, Panel D reveals that the asymmetry proxy sometimes moves into negative territory (*e.g.*, in early 2020). Hence, on these days, the asymmetry is reversed: the average reaction to extreme positive returns is stronger than the average reaction to extreme negative returns. Finally, in Panel E, we observe that the high speed of reaction of RH investors to past extreme negative returns, while valid for most of the period, can also disappear on some days (see *e.g.*, early days of July, 2018).

It is also interesting to note that our sample includes the COVID-19 pandemic announcement (March, 2020). A closer look at these daily series (*e.g.*, Panel A) suggests that the announcement might have triggered a sudden change in RH investors' behavior. For this reason, we devote a subsection to analyzing how the COVID-19 pandemic announcement might affect our results.

5.3. Time-series factors

We propose to test for three factors impacting RH investors' reactions to previous returns. First, we include Cboe Volatility Index (VIX) daily series from CRSP. We think this index which represents market volatility (and is often referred to as the fear index) is a key indicator that most retail investors pay attention to. We expect that RH investors' reactions to previous returns is negatively related to the level of the VIX. Indeed, in a high-volatility environment, it is natural to think that investors will be more cautious, therefore reacting less strongly to extreme price movements. Second, we test for the market excess returns daily series from Kenneth R. French's website. In our view, the daily market trend is also likely to influence the trading behavior of RH investors. Periods of rising markets might be favorable to stronger reaction to previous extreme price movements, and vice versa. We therefore expect a positive relationship between most of our proxies and this factor. Third, we incorporate the sentiment index series from the American Association of Individual Investors

(AAII) which measures the percentage of bullish investors among all the AAII members interviewed.[15] Since we believe that most of the RH investors' behaviors we measure are driven by non-rational factors, we expect a significant positive correlation between them and this sentiment index.

Panel A of Table 3 presents summary statistics for these factors. The VIX index has an average level of 20.21 during our sample period and is quite volatile, ranging from 10.85 to 82.69. The average daily market excess return is almost zero, at 0.03%. The investors sentiment index has an average of 33% bullish investors over the period, ranging from 20% to 46%.

5.4. Results

5.4.1. Main specification

Our main specification is as follows:

$$Y_d = \lambda_0 + \lambda_{VIX} VIX_d + \lambda_{MKT} MKT_d + \lambda_{SENT} SENT_d + \epsilon_d\,, \tag{10}$$

where Y_d is one the proxy variables defined in equations (5) through (9) and VIX_d, MKT_d and $SENT_d$ are the three time-series factors defined above.

Table 4 presents estimation results. The VIX coefficient is negative and significant at the 1% level for all dependent variables with the exception of $Asy^{(d)}$. This indicates that, on low-volatility days, RH investors' reaction to extreme returns tend to be stronger ($Ext^{(d)}$), and this is valid for both negative and positive extreme returns ($ExtNeg^{(d)}$ and $ExtPos^{(d)}$). The speed of reaction of RH investors to previous returns in the last hour is also higher on low-volatility days. The asymmetry of reaction to extreme returns, however, does not seem to be influenced by the level of market volatility. Overall, this result suggests that the level of market volatility tend to decrease the activity of RH investors. On volatile days, they react less strongly to past extreme returns, and more slowly to past extreme negative returns. On calm days, they react more strongly to past extreme returns, and more quickly to past extreme negative returns. Therefore, these results are inline with our expectations.

The MKT coefficient is not statistically different from zero for the $Ext^{(d)}$ dependent variable but interestingly, when we split the reaction into extreme negative and extreme positive returns, the estimates have opposite signs. It is positive for $ExtNeg^{(d)}$ (10% level) but negative for $ExtPos^{(d)}$ (1% level). This suggests that, on days of upside market move, RH investors tend to respond more strongly to extreme negative returns but less strongly to extreme positive returns. In other words, as the coefficient relative to the $Asy^{(d)}$ regression confirms, the asymmetry of reaction to previous extreme returns (toward negative returns) is stronger on days of market growth, and lower on days of market downturn. For

[15] The index is based on a survey asking AAII members the following question each week: "What direction do you feel the stock market will be in the next 6 months?" the answers are then classified into three categories: (i) bearish, (ii) neutral, (iii) bullish. The original data is at weekly frequency. We transform it into daily frequency by linear interpolation.

the $SpeedExtNeg^{(d)}$ dependent variable, the estimate is positive and significant (10% level). Therefore, on days when the market moves up, RH investors tend to respond more quickly to stocks exhibiting past extreme negative returns. As we expected a positive correlation between this variable and most of our RH investors' behavior proxies, the results here are only partially in line with our prior.

The $SENT$ coefficient is strongly positive and significant for the $Ext^{(d)}$ and $ExtNeg^{(d)}$ dependent variables. It is however not statistically different from zero for the $ExtPos^{(d)}$ dependent variable. Therefore, the positive correlation between RH investors' reaction to past extreme returns and investors sentiment mostly come from the negative side of the extreme returns: during days of optimism, RH investors tend to respond more strongly to large negative returns, but not necessarily to large positive returns. While this result might suggest that investors sentiment should affect the asymmetry of reaction to extreme returns, the coefficient is not significant when the $Asy^{(d)}$ is the dependent variable.[16] Finally, investors sentiment seems to affect, whereas weakly (10% level), the speed of reaction of RH investors' to extreme negative return: on days of positive sentiment, RH investors tend to respond faster to past negative returns.

5.4.2. Impact of COVID-19

To evaluate the sensitivity of these results with respect to the COVID-19 pandemic announcement, we introduce an indicator function taking the value of one if the observation belongs to the pre-COVID period, and zero if the observation belongs to the post-COVID period. As a cutoff date, we use the date when the World Health Organization declared the COVID-19 outbreak as a pandemic: March 11, 2020. We then multiply each factor of regressions (10) by this dummy variable, and test for the difference between the factor in the pre- and post-COVID announcement period.

Table 5 presents the results. The estimation results displayed in Panel A shows that splitting the effect of the factors into the pre- and post-COVID announcement period increases significantly the quality of the fit. Indeed, the R-squared of these regressions are significantly higher than the R-squared of the main specification (Table 4).

Panel B presents the results from F-tests evaluating coefficients' equality before and after the announcement date. Specifically, for each factor F, the null hypothesis is that $\hat{\beta}_{F.preCOV} = \hat{\beta}_{F.postCOV}$. The difference between the impact of market volatility before and after the announcement is significant at the 1% level for the $Ext^{(d)}$ dependent variable. In the pre-COVID period, the negative relation between the VIX and RH investors' reaction to previous extreme return is stronger than after the announcement. Therefore, a pre-COVID-low-volatility day would make RH investors open more positions in stocks exhibiting past extreme returns, compared to a low-volatility day after the announcement date. In other words, the effect we observe in the main specification – that RH investors react less strongly

[16] A potential explanation of the non-significance of this coefficient is that results might be sensitive to alternative definitions of the asymmetry of reaction's proxy.

to past extreme returns on more volatile days – is significantly more pronounced in the period preceding COVID-19 market correction of March 2020. In addition, results with respect to the dependent variable $ExtNeg^{(d)}$ tell us that market volatility was an important determinant of RH investors' reactions to past extreme negative returns before the COVID announcement, but not after. In terms of speed of reaction, we see that the effect we observe in the main specification – that RH investors react more slowly to past extreme negative returns on more volatile days – is significantly more pronounced in the period preceding the COVID announcement.

With regards to the $SENT$ factor, Panel A shows that the coefficients are significant (and positive) for the same dependent variables as for the main specification, namely $Ext^{(d)}$, $ExtNeg^{(d)}$ and $SpeedExtNeg^{(d)}$, but only in the pre-COVID period. Indeed, after the announcement, none of the $SENT$ coefficients are significantly different from zero. Therefore, it appears that the correlation between our RH investors' behavior proxies and investors sentiment disappear after the COVID-19 market correction.

Overall, these results show that the COVID-19 announcement tend to mitigate the effects of the time series factors on RH investors' behavior that we observed over the entire period. Indeed, market volatility has a more pronounced effect before the announcement than after, and the investors sentiment effect observed over the entire period is only significant in the pre-COVID period.

6. Cross sectional variation in RH investors' behavior and its determinants

6.1. Reaction of RH investors to high-frequency returns by company

We follow the same steps as in the previous section to identify the determinants of RH investors' reaction to previous returns in the cross-section. Before running regression (4) separately for each company, we adjust our sample based on two criteria. First each company must have at least one year of observations. Second, the company must be an "active Robinhood stock", in the sense that its associated RH users change variable ($\Delta N_{i,t_{i,k}}$) must be non-zero for at least 50% of observations. After these adjustments, our sample reduces to 1,454 companies. With 1,454 companies and five horizons, we estimate a total of 7,270 regressions. We denote the coefficients of interest representing the average reaction of RH investors to a given group of return level g at a given horizon h and a given company i as $\beta_{g,h}^{(i)}$. The average number of observations per regression is 3,089. Figure 3 displays the distribution of the $\beta_{g,h}^{(i)}$ for all groups of return levels at each horizon. The distributions have strong dispersion. For example, as Panel A shows, for some companies, RH investors' react to stocks exhibiting extreme negative returns in the past few hours by closing their positions in that stock ($\beta_{1,1}^{(i)}$ can be as low as -10). For other companies however, they respond by opening new positions in that stock ($\beta_{1,1}^{(i)}$ can be as high as 30). This heterogeneity across firms seems to be present for all coefficients and horizons. Therefore, the three RH investors' behavioral patterns described earlier must also vary across firms.[17]

6.2. Proxy variables and firm characteristics

First, we construct variables analogous to equations (5) through (9), except that they are now based on the regressions for each company. In other words, we replace the $\hat{\beta}_{g,h}^{(d)}$ from these equations by $\hat{\beta}_{g,h}^{(i)}$ and obtain the five proxies $Ext^{(i)}$, $ExtNeg^{(i)}$, $ExtPos^{(i)}$, $Asy^{(i)}$ and $SpeedExtNeg^{(i)}$.

We construct various firm characteristics using data from COMPUSTAT and CRSP.[18] that are defined in Table 6. We believe that these characteristics may influence how RH investors respond to previous returns for a given firm. For example, it is possible that RH investors focus their attention on bigger or smaller firms (*SIZE*). They may also adopt different behaviors depending on the investing style (value vs growth) or the firm's financial health. Therefore we include the Book-to-Market (*BM*) and Debt-to-Asset (*DA*) ratios. The other characteristics are market-related. Among them, we use trading volume (*VOL*), liquidity level (*BIDASK*), firm's systematic risk (*BETA*) and idiosyncratic risk (*IVOL*). Because RH investors' behavior could depend on the industry of the firm, we also obtain information on firms' sectors, as per the General Industrial Classification Standard (GICS).

Panel B of Table 3 presents summary statistics for firm characteristics. Among other characteristics, the average (log) size of the 1,426 firms in our sample is 21.20 (approximately $600 millions). The average indebtedness as measured per the Debt-to-Asset ratio is about 32%. The median firm systematic risk, as measured by the market beta, is 1.18. In terms of specific risk, measured by the standard deviation of estimated residuals in the Fama and French (1993) 3-factors model, firms range from 0.81% to 27.84%.

6.3. Results

6.3.1. Heterogeneity at the industry-level

As a first look, we estimate the average of our five proxies for each industry. Figure 6 shows that the various RH investors' behavioral patterns with respect to previous returns are heterogeneous across industry. For example, RH investors tend to respond stronger to firms exhibiting extreme returns ($Ext^{(i)}$) from the Energy sector (average RH users change of 8.38BP) but weaker to those from the Utility sector (3.25BP). The asymmetry of reaction to extreme returns is more pronounced for firms in the Finance sector, and particularly weak for firms in the Health Care sector. The distribution of the RH investors' speed of reaction to large negative returns across industry seems less dispersed, as all conditional averages are between 4.26BP and 6.64BP.[19]

[17] For more details on the estimation results of the company regressions, refer to section A.5 of the appendix.

[18] data fields used from COMPUSTAT: *MKVALTQ*, *CEQQ*, *DLCQ*, *DLTTQ*, *ATQ*, *GSECTOR*. Data fields used from CRSP: *vol*, *bid*, *ask*.

[19] These averages are estimated within a regression setup. In addition, the results related to the Utility and Real Estate sectors must be taken with caution as only 9 and 31 companies of our sample belong to these respective sectors. For more details on the estimation results and the number of firms in each sector, refer to Section A.5 of the appendix.

6.3.2. Main specification

Our main specification is as follows

$$
\begin{aligned}
Y_i = {} & \lambda_0 + \lambda_{SIZE} SIZE_i + \lambda_{BM} BM_i + \lambda_{DA} DA_i + \lambda_{VOL} VOL_i + \lambda_{BIDASK} BIDASK_i \\
& + \lambda_{BETA} BETA_i + \lambda_{IVOL} IVOL_i + \sum_s \lambda_{IND,s} IND_{i,s} + \epsilon_i \,,
\end{aligned} \tag{11}
$$

where Y_i is one the proxy variables defined earlier, and $IND_{i,s}$ is a categorical variable taking a value of one if firm i belongs to sector s, and zero otherwise. The other regressors are firm characteristics.

Table 7 presents estimation results. The $SIZE$ coefficient is not statistically different from zero for the regression with $Ext^{(i)}$ as dependent variable. However, splitting the RH investors' reaction to extreme returns by returns' sign reveals that the estimate is actually positive and significant for negative extreme returns, but negative and significant for positive extreme returns (on aggregate, these two effects cancel out). Therefore, when larger firms exhibit extreme negative returns, RH investors tend to respond more strongly by opening more position in these firms. Conversely, when larger firms exhibit extreme positive returns, RH investors tend to respond more weakly by opening fewer position in these firms. It follows that the asymmetry of reaction to extreme returns is stronger for bigger firms, as emphasized by the positive and significant coefficient for the specification with $Asy^{(i)}$ as dependent variable. The size of the firm also seems to influence the speed of reaction of RH investors to previous large negative returns, with a positive and significant (1% level) coefficient. Therefore, the bigger the firm, the faster RH investors respond to an extreme negative return.

The Book-to-market ratio of the firm has a moderate influence (5% level) on RH investors' reaction to extreme negative returns ($ExtNeg^{(i)}$) and the asymmetry of their reaction to extreme returns ($Asy^{(i)}$). As the estimate is positive in both cases, RH investors tend to respond more strongly to "value firms" that exhibit large negative price movements in the previous hour. Moreover, the higher book-to-market ratio of the firm (*i.e.* the more "value" the firm is), the stronger is the asymmetry of reaction of RH investors to previous extreme returns toward negative returns.

The level of indebtedness of the firm seems to be an important determinant for most of the cross-section variations of our results. The DA coefficient is positive and significant at the 1% level for all dependent variables with the exception of $ExtPos^{(i)}$. RH investors tend to react more strongly to extreme returns for firms with a lot of debt relative to their asset size. For these firms, RH investors also tend to have a more asymmetric response to extreme returns, and their reaction speed to extreme negative returns is higher.

With regards to trading volume of the firm, RH investors tend to react more strongly to higher-volume stocks exhibiting extreme returns. However, splitting the reaction to extreme returns by return's sign reveals that the effect is only significant for the extreme positive returns. It follows that the coefficient on regression with the $Asy^{(i)}$ dependent variable is negative (1% level). It means that, the higher the trading volume of the firm, the lower is the asymmetry of reaction of RH investors to extreme returns on that firm. Therefore, for these high-volume firms, the "smirk" we observed earlier may be reversed (*i.e.*, the higher point being on the positive – not negative – side).

The level of illiquidity of the firm, as measured by its bid-ask spread, is positively related to the RH investors' reaction to extreme returns of that firm in the last hour (1% level). The positive relation is stronger for the reaction to extreme positive returns than for the reaction to extreme negative returns. It follows that the $BIDASK$ coefficient is negative (and significant, 5% level) for the $Asy^{(i)}$ dependent variable. Therefore, for the most illiquid firms, the asymmetry of reaction of RH investors to extreme returns may be reversed.

Systematic risk seems to be an important determinant. The $BETA$ coefficients are positive and significant for all but the $ExtPos^{(i)}$ dependent variables. The higher the volatility of the firm relative to the overall market, the stronger the reaction of RH investors to extreme returns on that firm. This stronger reaction is actually driven by the extreme negative returns, as shown by the high and positive coefficient for the $ExtNeg^{(i)}$ specification. Therefore, as confirmed by the coefficient with the $Asy^{(i)}$ specification, RH investors' reaction to extreme return is more asymmetric for higher-beta firms. The estimate is also positive and significant for the $SpeedExtNeg^{(i)}$ dependent variable, meaning that RH investors tend to respond faster to higher-beta firms exhibiting large past negative price movements.

Idiosyncratic risk also matters. For all dependent variables, the coefficients are positive and significant at the 1% level. Therefore, for firms with high specific risks, RH investors tend to respond more strongly to previous extreme return, their asymmetry of reaction to extreme returns is more pronounced, and their response to large negative price movements is quicker.

6.3.3. Impact of COVID-19

This section investigates the effect on the COVID-19 pandemic announcement on the cross-sectional variations of the key RH investors behavioral patterns observed earlier. In order to compare the results in the pre- and post-COVID announcement period, we first estimate regression (4) for each company and each sub period. This allows us to determine the average reaction of RH investors to a given group of return level g at a given horizon h and a given company i in the respective sub period, and compute the proxy variables analogous to equations (5) to (9).[20]

Industry-level. For each industry, we estimate the average of our five proxies per sub period. With eleven GICS sectors and two sub-periods, we obtain a total of 22 estimates.[21] Figure 7 displays the difference between the averages of each proxy per sector before and after the COVID-19 pandemic announcement. The average reactions per sector to extreme returns

[20]Distributions of these proxies are shown in section A.6 of the appendix.

[21]These averages are estimated within a regression setup. For more detail on the estimation results, refer to Section A.6 of the appendix.

($Ext^{(i)}$) before and after the announcement are not significantly different (5% level), except for the Health Care sector. For firms belonging to this sector, RH investors tended to respond less strongly to extreme returns after the COVID-19 market correction of March, 2020. The average reactions per sector to extreme negative returns ($ExtNeg^{(i)}$), however, seem to vary more from one period to the other. For the Energy, Consumer Discretionary, Health Care and Information Technology sectors, RH investors tended respond more strongly to extreme negative returns in the pre-COVID announcement period. For the Utility sector however, they tended to respond more strongly after the market correction. Panel D reveals that, for the Consumer Discretionary, Health Care, Energy and Information Technology sectors, the asymmetry of reaction to extreme returns tended to be less pronounced after the announcement. Therefore, RH investors tended to respond to extreme return from companies in these sectors in a more symmetric way after the COVID-19 market correction of March 2020 (*i.e.*, from March to August 2020, they tended to adopt a similar attitude toward extreme negative and extreme positive returns). Moreover, Panel E shows that after the COVID-19 market correction, RH investors slowed down in their reaction to extreme negative returns from companies in the Energy, Consumer Discretionary, Health Care, Information Technology and Telecommunications sectors.[22]

Company-level. We construct a variation of regression (11) where each independent variable is multiplied by a dummy variable taking the value of one if the Y_i observation belong to the pre-COVID period, and zero otherwise.[23] Table 8 presents estimation results. In Panel A, each coefficient represents the effect of the characteristic before and after the COVID-19 pandemic announcement, after controlling for the (pre- and post-COVID announcement) effect of the industry. Compared to regression (11) which does not differentiate for the pre- and post-COVID announcement period, this specification returns higher R^2s for most proxies. For example, the R^2 for the specification evaluating the reaction to extreme return and the reaction to extreme negative returns are 56% and 61%, compared to 25% and 23% in Table 7, respectively. Similarly, regarding the speed of reaction to extreme negative returns, the R^2 is 31% compared to only 4% in Table 7. This suggests that for these three types of behavior, splitting the effect of the firm characteristics and industries before and after the COVID-19 pandemic announcement improves significantly the explanation power. This improvement however, does not manifest for the two other proxies, $ExtPos^{(i)}$ and $Asy^{(i)}$.

To compare formally the effect of each characteristic before and after the announcement, we test for coefficients' equality. Panel B displays the p-values associated to these tests. For the $Ext^{(i)}$ proxy, none of the coefficients are significantly different in the two sub periods, except the one associated to the illiquidity ($BIDASK$) characteristic (10% level). This result suggests that after the COVID-19 pandemic announcement, RH investors tended to respond

[22]Note that these results must be nuanced by the fact that the number of firms per sector is uneven. For example, as section A.5 shows, the Utility or Real Estate sectors contain only 31 and 9 firms, respectively.

[23]For an explicit description of the specification, refer to section A.6 of the appendix.

more strongly to more illiquid firms exhibiting past extreme returns. Before the announcement, this phenomena was also existing, but at a lower magnitude. For the $ExtNeg^{(i)}$ proxy, none of the coefficients are significantly different. This suggests that in general, the effects of the various firm characteristics on RH investors reaction to extreme negative returns that we observed in the full period are valid both before and after the announcement. For the reaction to extreme positive returns ($ExtPos^{(i)}$), however, the effect of trading volume, illiquidity and specific risk of the firm seem to differ pre- and post-COVID announcement. Therefore, after March 11, 2020, RH investors tended to respond more strongly to high-volume stocks exhibiting past extreme positive returns than before this date. Similarly, they tended to respond more strongly to illiquid stocks exhibiting past extreme positive returns. Interestingly, the coefficient relative to specific risk is not different from zero in the pre-COVID period, but significantly positive post-announcement. It means that the level of idiosyncratic risk of the firm had a positive influence on how RH investors react to extreme positive returns after the COVID-19 pandemic announcement, but not before. Most of the coefficients do not differ in the pre- and post-COVID period for the asymmetry of reaction to past extreme returns proxy ($Asy^{(i)}$). The specific risk coefficient is however significantly positive in the pre-COVID sub period, but not different from zero in the post-COVID sub period. Hence, the level of idiosyncratic risk of the firm tended to increase the asymmetry of reaction to extreme returns before the COVID-19 pandemic announcement, but not after. In terms of speed of reaction to extreme negative returns ($SpeedExtNeg^{(i)}$), the $SIZE$ estimate is positive and significant in the post-COVID sub period, but not significant in the pre-COVID. The difference is significant at the 10% level. This suggests that after the COVID market correction of March 2020, RH investors responded more quickly to big-capitalization firms experiencing large negative price movements. Before the announcement however, no such relation seems to exist. The systematic risk coefficient is positive and significant in the pre-COVID sub period but not significantly different from zero after the announcement. Therefore, before the announcement, RH investors tended to respond more quickly to higher-betas firms, but it was no longer the case after the announcement.

7. Conclusion

We identify three key intraday behaviors of Robinhood users. First, these investors tend to open new positions in stocks that have experienced extreme returns in the last few hours. Second, the magnitude of their reaction is larger (smaller) for stocks exhibiting past extreme negative (positive) returns. Finally, they tend to be particularly quick in their reaction to stocks that have experienced large negative price movements.

Building on these findings, we attempt to identify their determinants in the time series and cross section. Market volatility appears to be an important determinant in the time series: On low-volatility days, Robinhood investors tend to respond more strongly to extreme returns and more quickly to extreme negative returns. Most of the behaviors we study are also positively correlated with investors sentiment: On positive-sentiment periods, they exhibit stronger reactions on stocks exhibiting past extreme returns. They also tend to respond more quickly to stocks exhibiting extreme negative returns. For cross-sectional variations, we find that for larger firms, Robinhood investors respond more strongly when they exhibit extreme negative returns but less strongly when they exhibit extreme positive returns. They also tend to have a higher speed of reaction for larger firms exhibiting past negative returns. In addition, most of our Robinhood investors proxies are positively correlated to characteristics that increase the uncertainty around the firm. For example, investors tend to react more strongly to stocks with high leverage, low liquidity levels, high systematic and idiosyncratic risks. The speed of reaction is also higher for stocks with these characteristics.

Finally, the market correction associated to the COVID-19 pademic announcement appears to have a non-negligible effect on the intraday behavior of Robinhood investors. For instance, the impact of market volatility on reaction to extreme returns is significantly more pronounced before the pandemic announcement in March 2020. The effect of investors' sentiment is important in the pre-COVID period, but disappears after. For cross-sectional variations, firms characteristics such as idiosyncratic risk appear to be important in determining the asymmetry of reaction to extreme returns before the COVID announcement, but not after.

References

Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. Review of Financial Studies 28, 33–72. doi:10.1093/rfs/hhu059.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645–1680. doi:10.1111/j.1540-6261.2006.00885.x.

Barber, B.M., Huang, X., Odean, T., Schwarz, C., 2021. Attention induced trading returns: Evidence from Robinhood users. Working paper doi:10.2139/ssrn.3715077.

Barber, B.M., Lee, Y.T., Liu, Y.J., Odean, T., 2009a. Just how much do individual investors lose by trading? The Review of Financial Studies 22, 609–632.

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, 785–818. doi:10.1093/rfs/hhm079.

Barber, B.M., Odean, T., Zhu, N., 2009b. Do retail trades move markets? Review of Financial Studies 22, 151–186. doi:10.1093/rfs/hhn035.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: trades and quotes. Econometrics Journal 12, C1–C32. doi:10.2307/23116045.

Barron's, 2021. Robinhood is going public. 4 metrics that could determine its value. URL: https://www.barrons.com/articles/robinhood-ipo-stock-value-51625166659. *By Avi Salzman, last updated July 2, 2021.*

Ben-David, I., Franzoni, F., Kim, B., Moussawi, R., 2021. Competition for attention in the ETF space. Working paper doi:10.3386/w28369.

Black, F., 1986. Noise. Journal of Finance 41, 529–543. doi:10.2307/2328481.

Boehmer, E., Jones, C.M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. Journal of Finance 76, 2249–2305. doi:https://doi.org/10.1111/jofi.13033.

BusinessofApps, 2021. Robinhood revenue and usage statistics (2021). URL: https://www.businessofapps.com/data/robinhood-statistics/. *By David Curry, last updated July 11, 2021.*

Chapkovski, P., Khapko, M., Zoican, M., 2021. Does gamified trading stimulate risk taking? doi:10.2139/ssrn.3971868. Working paper.

CNBC Markets, 2020. Young investors pile into stocks, seeing 'generational-buying moment' instead of risk. URL: https://www.cnbc.com/2020/05/12/young-investors-pile-into-stocks-seeing-generational-buying-moment-instead-of-risk.html. *By Maggie Fitzgerald, last updated May 12, 2020.*

CNBC Markets, 2021. Robinhood gets rid of confetti feature amid scrutiny over gamification of investing. URL: https://www.cnbc.com/2021/03/31/robinhood-gets-rid-of-confetti-feature-amid-scrutiny-over-gamification.html. *By Maggie Fitzgerald, last updated March 31, 2021.*

CNN Business, 2020. Regulators slam Robinhood for targeting newbie investors with aggressive marketing. URL: https://www.cnn.com/2020/12/16/business/robinhood-trading-app-complaint-massachusetts/index.html. *By Matt Egan, last updated December 16, 2020.*

Cookson, J.A., Niessner, M., 2020. Why don't we agree? Evidence from a social network of investors. Journal of Finance 75, 173–228. doi:10.1111/jofi.12852.

Coval, J.D., Hirshleifer, D., Shumway, T., 2021. Can individual investors beat the market? Review of Asset Pricing Studies 11, 552–579. doi:10.1093/rapstu/raab017.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. Journal of Finance 66, 1461–1499. doi:10.1111/j.1540-6261.2011.01679.x.

Eaton, G.W., Green, T.C., Roseman, B., Wu, Y., 2021. Zero-commission individual investors, high frequency traders, and stock market quality. Working paper doi:10.2139/ssrn.3776874.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56. doi:0.1016/0304-405X(93)90023-5.

Fong, K.Y.L., Gallagher, D.R., Lee, A.D., 2014. Individual investors and broker types. Journal of Financial and Quantitative Analysis 49, 431–451. doi:10.1017/S0022109014000349.

Forbes Advisor, 2021. Should you invest in the Robinhood IPO? URL: https://www.forbes.com/advisor/investing/robinhood-ipo/. *By Taylor Tepper and Benjamin Curry, last updated July 30, 2021.*

Fortune, 2020. Robinhood will no longer share stock 'popularity data' with sites like robintrack. URL: https://fortune.com/2020/08/10/robinhood-popularity-data-robintrack-stock-market-trading-tracker/. *By Jeff John Roberts, last updated August 10, 2020.*

Friedman, H.L., Zeng, Z., 2021. Retail investor trading and market reactions to earnings announcements. doi:10.2139/ssrn.3817979. Working paper.

Gargano, A., Rossi, A.G., 2018. Does It Pay to Pay Attention? Review of Financial Studies 31, 4595–4649. doi:10.1093/rfs/hhy050.

Hvidkjaer, S., 2008. Small trades and the cross-section of stock returns. Review of Financial Studies 21, 1123–1151. doi:10.1093/rfs/hhn049.

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, 337–356. doi:10.1016/j.jfineco.2016.11.003.

Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. Journal of Finance 63, 273–310. doi:10.1111/j.1540-6261.2008.01316.x.

Kelley, E.K., Tetlock, P.C., 2013. How wise are crowds? insights from retail orders and stock returns. Journal of Finance 68, 1229–1265. doi:10.1111/jofi.12028.

Kumar, A., Lee, C.M.C., 2006. Retail investor sentiment and return comovements. Journal of Finance 61, 2451–2486. doi:10.1111/j.1540-6261.2006.01063.x.

Lou, D., Polk, C., Skouras, S., 2019. A tug of war: Overnight versus intraday expected returns. Journal of Financial Economics 134, 192–213. doi:10.1016/j.jfineco.2019.03.011.

Moss, A., Naughton, J.P., Wang, C., 2020. The irrelevance of ESG disclosure to retail investors: Evidence from Robinhood. Working paper doi:10.2139/ssrn.3604847.

New York Times, 2020. Robinhood has lured young traders, sometimes with devastating results. URL: https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html. *By Nathaniel Popper, last updated July 21, 2021.*

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708. doi:10.2307/1913610.

Ozik, G., Sadka, R., Shen, S., 2021. Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. Journal of Finance and Quantitative Analysis (forthcoming). doi:10.2139/ssrn.3663970.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. The Journal of Finance 40, 777–790. doi:10.2307/2327802.

Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288–302. doi:10.1016/j.jfineco.2011.12.001.

Statista, 2021. Number of users of robinhood from 2014 to 3rd quarter 2021. URL: https://www.statista.com/statistics/822176/number-of-users-robinhood/. *By Fanny Norrestad, last updated December 13, 2021.*

The Economist, 2021. Just how mighty are active retail traders? URL: https://www.economist.com/finance-and-economics/2021/08/21/just-how-mighty-are-active-retail-traders. finance & Economics, August 21st 2021 edition.

Van der Beck, P., Jaunin, C., 2021. The equity market implications of the retail investment boom. Swiss Finance Institute Research Paper Series doi:10.2139/ssrn.3696066.

Welch, I., 2021. The wisdom of the Robinhood crowd. Journal of Finance (forthcoming). doi:10.2139/ssrn.3696066.

Yuan, Y., 2015. Market-wide attention, trading, and stock returns. Journal of Financial Economics 116, 548–564. doi:10.1016/j.jfineco.2015.03.006.

Table 1: Summary statistics of RH users change and high-frequency adjusted returns
This table presents summary statistics across stock-day-time observations for our measure of Robinhood users' change ($\Delta N_{i,t_{i,k}}$) and high-frequency adjusted returns ($r_{i,t_{i,k}}$). Statistics associated to $\Delta N_{i,t_{i,k}}$ are expressed in basis point, except for *Nobs*, T and #, which represent the number of observations, trading days, and companies, respectively.

	Av	Std	Min	25th	50th	75th	Max	*Nobs*	T	#
$\Delta N_{i,t_{i,k}}$	4.36	115.28	-15126.30	-3.05	0.00	3.31	59,602.68	8,031,495	527	2,853
$r_{i,t_{i,k}}$	-0.01	1.00	-30.33	-0.41	0.00	0.41	44.42	8,031,495	527	2,853

Table 2: Regression results

This table shows the estimates of interest, $\beta_{g,h}$, from regression (4). All five regressions are based on the full panel (all stock-day-hour observations) and are estimated by ordinary least squares. Estimates are expressed in basis point. Associated t-statistics are shown in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
<-2	7.95***	6.07***	4.98***	4.2***	3.57***
	(148.44)	(112.97)	(92.75)	(78.16)	(66.54)
[-2,-1[	2.54***	2.1***	1.62***	1.29***	0.99***
	(83.69)	(69.15)	(53.23)	(42.5)	(32.63)
[-1,0[	0.31***	0.28***	0.21***	0.17***	0.15***
	(24.13)	(21.74)	(16.32)	(13.18)	(11.69)
[0,1[	-0.2***	-0.16***	-0.05***	0.02	0.08***
	(-16.73)	(-12.96)	(-3.78)	(1.55)	(6.63)
[1,2[	0.03	0.28***	0.47***	0.62***	0.68***
	(0.84)	(9.21)	(15.3)	(20.32)	(22.25)
>2	2.71***	2.61***	2.68***	2.7***	2.78***
	(49.59)	(47.74)	(48.94)	(49.33)	(51.04)
R^2	0.014	0.015	0.014	0.014	0.014
$Nobs$	8,017,230	8,017,230	8,017,230	8,017,230	8,017,230

Table 3: Summary Statistics of time-series factors and firm characteristics
Panel A presents summary statistics across days for the time-series factors. VIX is the level of the Cboe Volatility Index, MKT is the excess market return over the riskless rate, expressed in percent and $SENT$ is the percentage of bullish investors as measured by the AAII. Panel B presents summary statistics across firms for the characteristic values defined in Table 6.

Panel A: Time-series factors								
	Av	Std	Min	25th	50th	75th	Max	T
VIX	20.21	10.95	10.85	13.40	15.96	22.65	82.69	526
MKT	0.03	1.60	-12.00	-0.41	0.12	0.69	9.34	526
SENT	33.13	5.63	20.23	29.53	33.87	36.80	45.66	526
Panel B: Firm characteristics								
	Av	Std	Min	25th	50th	75th	Max	N
SIZE	21.20	2.38	15.22	19.51	21.17	23.00	27.65	1,426
BM	0.44	1.08	-19.74	0.13	0.30	0.63	11.81	1,426
DA	0.32	0.30	0.00	0.10	0.28	0.44	5.40	1,426
VOL	13.65	1.31	9.59	12.68	13.66	14.51	18.11	1,426
BIDASK(%)	0.32	0.50	0.01	0.04	0.10	0.34	3.98	1,426
BETA	1.19	0.44	-0.73	0.92	1.18	1.45	3.12	1,426
IVOL(%)	3.98	2.63	0.81	2.09	3.34	5.05	27.84	1,426

Table 4: Time-series regression results

Each column corresponds to the proxy variable defined in equations (5) to (8), respectively. All specifications are estimated by standard OLS. The associated t-statistics in parenthesis have been computed with Newey and West (1987) standard errors that are robust to heteroskedasticity and autocorrelation up to 22 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	$Ext^{(d)}$	$ExtNeg^{(d)}$	$ExtPos^{(d)}$	$Asy^{(d)}$	$SpeedExtNeg^{(d)}$
Intercept	4.33***	4.64***	4.02*	0.62	5.70***
	(3.18)	(3.26)	(1.71)	(0.23)	(3.65)
VIX	-0.09***	-0.08***	-0.10***	0.02	-0.08***
	(-6.22)	(-3.74)	(-4.31)	(0.59)	(-5.12)
MKT	-0.04	0.31*	-0.40***	0.71**	0.34*
	(-0.72)	(1.71)	(-2.63)	(2.27)	(1.89)
SENT	12.44***	18.56***	6.33	12.23	8.49*
	(2.93)	(4.04)	(0.80)	(1.25)	(1.75)
Adj.R^2	0.175	0.133	0.078	0.034	0.071
Nobs	526	526	526	526	526

Table 5: Impact of the COVID-19 pandemic on time-series regression results
Each column corresponds to the proxy variable defined in equations (5) to (8), respectively. Panel A present estimation results using standard OLS method. The associated t-statistics in parenthesis have been computed with Newey and West (1987) standard errors that are robust to heteroskedasticity and autocorrelation up to 22 lags. Panel B presents the significance of the tests on coefficient equality between the pre- and post-COVID period. For each factor F, the null hypothesis is that $\hat{\beta}_{F.preCOV} = \hat{\beta}_{F.postCOV}$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Panel A: estimation results					
	$Ext^{(d)}$	$ExtNeg^{(d)}$	$ExtPos^{(d)}$	$Asy^{(d)}$	$SpeedExtNeg^{(d)}$
Int.preCOV	5.78**	7.65***	3.91	3.75	6.90***
	(2.35)	(2.79)	(1.18)	(1.05)	(2.94)
Int.postCOV	7.68***	8.15***	7.22**	0.93	11.16***
	(7.73)	(4.38)	(2.18)	(0.19)	(9.44)
VIX.preCOV	-0.21***	-0.27***	-0.14*	-0.13*	-0.22***
	(-2.77)	(-3.10)	(-1.87)	(-1.85)	(-3.97)
VIX.postCOV	-0.07***	-0.03	-0.10**	0.07	-0.06***
	(-4.02)	(-1.16)	(-2.41)	(1.12)	(-2.81)
MKT.preCOV	-0.11	0.39**	-0.61***	1.00***	0.44**
	(-1.10)	(2.08)	(-3.19)	(3.14)	(2.07)
MKT.postCOV	-0.13	0.05	-0.32	0.37	0.07
	(-1.57)	(0.67)	(-1.50)	(1.36)	(0.80)
SENT.preCOV	13.33***	18.27***	8.39	9.87	11.18**
	(2.6)	(3.23)	(0.93)	(0.90)	(2.01)
SENT.postCOV	-0.14	3.40	-3.68	7.08	-8.2
	(-0.03)	(0.58)	(-0.32)	(0.44)	(-1.57)
Adj.R^2	0.850	0.844	0.480	0.421	0.731
Nobs	526	526	526	526	526

Panel B: tests of coefficients' equality					
	$Ext^{(d)}$	$ExtNeg^{(d)}$	$ExtPos^{(d)}$	$Asy^{(d)}$	$SpeedExtNeg^{(d)}$
Intercept					
VIX	***	***		**	***
MKT				*	
SENT	*				*

Table 6: Firm characteristics

This table outlines the firm characteristic variables used in the analysis. Variables $SIZE$, BM, DA are based on quarterly frequency data. Variables VOL, $BIDASK$, $BETA$ and $IVOL$ are based on daily frequency data. Data fields used from COMPUSTAT: *MKVALTQ*, *CEQQ*, *DLCQ*, *DLTTQ*, *ATQ*, *GSECTOR*. Data fields used from CRSP: *vol*, *bid*, *ask*.

Name	Code	Definition	Source
Size	SIZE	$\frac{1}{N_Q}\sum_{q=1}^{N_Q} log(MKTCAP_q)$	COMPUSTAT
Book-to-Market	BM	$\frac{1}{N_Q}\sum_{q=1}^{N_Q} \frac{\text{BV}_q}{MKTCAP_q}$	COMPUSTAT
Debt-to-Asset	DA	$\frac{1}{N_Q}\sum_{q=1}^{N_Q} \frac{\text{ST Debt}_q+\text{LT Debt}_q}{\text{Total Asset}_q}$	COMPUSTAT
Volume	VOL	$\frac{1}{N_D}\sum_{d=1}^{N_D} log(volume_d)$	CRSP
Bid-Ask spread	BIDASK	$\frac{1}{N_D}\sum_{d=1}^{N_D} \frac{ASK_d+BID_d}{MIDPOINT_d}$	CRSP
Systematic Risk	BETA	$\hat{\beta}$ from CAPM: $R_d = \alpha + \beta MKT_d + \epsilon_d$	CRSP, FRENCH's website
Idiosyncratic Volatility	IVOL	Std.Dev. of $\hat{\epsilon_d}$ from FF-3 factors: $R_d = \alpha + \beta MKT_d + \theta SMB_d + \phi HML_d + \epsilon_d$	CRSP, FRENCH's website
Sectors	SEC	Dummy variable corresponding to one of the 11 GICS sectors: *Energy (EN), Materials (MAT), Industrials (IND), Consumer Discretionary (CD), Consumer Staple (CS), Health Care (HLTH), Financials (FIN), Information Technology (IT), Utilities (UT), Real Estate (RE)*	COMPUSTAT

Table 7: Cross-section regression results

Each column corresponds to the proxy variable defined in section 6.2. While the associated coefficients are not reported, the regression includes the industry categorical variables, as specified in (11). The regressions are estimated by standard OLS. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	$Ext^{(i)}$	$ExtNeg^{(i)}$	$ExtPos^{(i)}$	$Asy^{(i)}$	$SpeedExtNeg^{(i)}$
Intercept	-11.69***	-30.49***	7.10***	-37.59***	-7.51***
	(-6.26)	(-11.89)	(2.77)	(-10.69)	(-3.46)
SIZE	0.04	1.29***	-1.21***	2.50***	0.41***
	(0.46)	(10.05)	(-9.37)	(14.16)	(3.77)
BM	0.10	0.33**	-0.13	0.47**	0.10
	(0.96)	(2.36)	(-0.96)	(2.42)	(0.87)
DA	1.17***	2.65***	-0.31	2.96***	1.26***
	(3.22)	(5.31)	(-0.62)	(4.32)	(2.99)
VOL	0.81***	0.24	1.39***	-1.14***	0.06
	(6.60)	(1.44)	(8.16)	(-4.91)	(0.40)
BIDASK(%)	2.90***	1.50***	4.31***	-2.81***	0.47
	(8.84)	(3.32)	(9.54)	(-4.53)	(1.24)
BETA	1.82***	4.06***	-0.43	4.49***	1.21***
	(7.06)	(11.49)	(-1.22)	(9.26)	(4.03)
IVOL(%)	0.55***	0.67***	0.43***	0.23**	0.35***
	(9.07)	(8.01)	(5.20)	(2.04)	(4.94)
Adj.R^2	0.247	0.227	0.412	0.407	0.042
Nobs	1,426	1,426	1,426	1,426	1,426

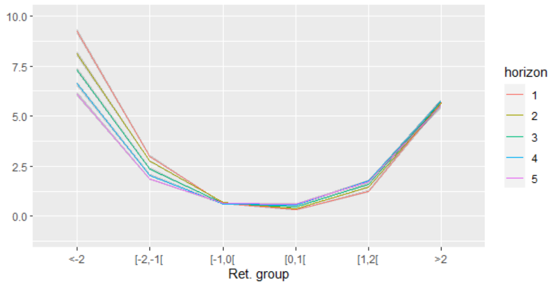
Table 8: Impact of the COVID-19 pandemic on cross-section regression results
Each column corresponds to the proxy variable defined in equations (5) to (8), respectively. Panel A present estimation results using standard OLS method. While the associated coefficients are not reported, the regression includes the industry categorical variables pre- and post-COVID, as specified in section A.6 of the appendix. Panel B presents the significance of the tests on coefficient equality between the pre- and post-COVID period. For each firm characteristics C, the null hypothesis is that $\hat{\beta}_{C.preCOV} = \hat{\beta}_{C.postCOV}$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Panel A: estimation results	$Ext^{(d)}$	$ExtNeg^{(d)}$	$ExtPos^{(d)}$	$Asy^{(d)}$	$SpeedExtNeg^{(d)}$
Int.preCOV	0.32	-20.24***	20.87***	-41.11***	-3.95
Int.postCOV	-9.53***	-19.51***	0.45	-19.95***	-6.16**
SIZE.preCOV	-0.33**	0.95***	-1.61***	2.56***	0.24
SIZE.postCOV	-0.18**	0.95***	-1.32***	2.27***	0.65***
BM.preCOV	-0.08	0.21	-0.37	0.57	-0.06
BM.postCOV	-0.02	-0.01	-0.04	0.02	-0.04
DA.preCOV	0.78	1.93***	-0.36	2.29**	1.71**
DA.postCOV	0.17	1.34**	-1.00*	2.34***	0.28
VOL.preCOV	0.83***	0.26	1.40***	-1.14***	0.12
VOL.postCOV	1.16***	0.27*	2.05***	-1.78***	-0.31*
BIDASK(%).preCOV	1.14**	0.02	2.27***	-2.25**	0.18
BIDASK(%).postCOV	2.11***	0.56	3.67***	-3.11***	0.23
BETA.preCOV	0.70*	3.15***	-1.76***	4.91***	1.93***
BETA.postCOV	0.11	2.27***	-2.04***	4.31***	0.30
IVOL(%).preCOV	0.43***	0.73***	0.13	0.61***	0.31***
IVOL(%).postCOV	0.48***	0.51***	0.45***	0.06	0.29***
Adj.R^2	0.559	0.609	0.347	0.444	0.310
Nobs	3,303	3,303	3,303	3,303	3,303

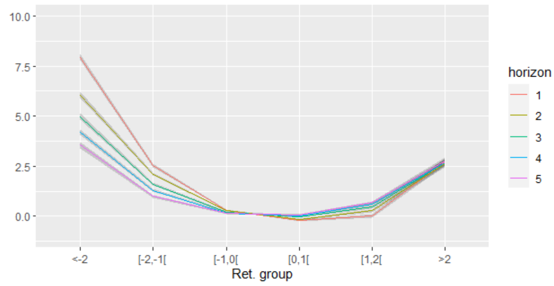
Panel B: tests of coefficients' equality	$Ext^{(d)}$	$ExtNeg^{(d)}$	$ExtPos^{(d)}$	$Asy^{(d)}$	$SpeedExtNeg^{(d)}$
Intercept	***		***	***	
SIZE					*
BM					
DA					
VOL			**		
BIDASK(%)	*		*		
BETA					**
IVOL(%)			**	***	

Figure 1: RH investors' reactions to stock returns

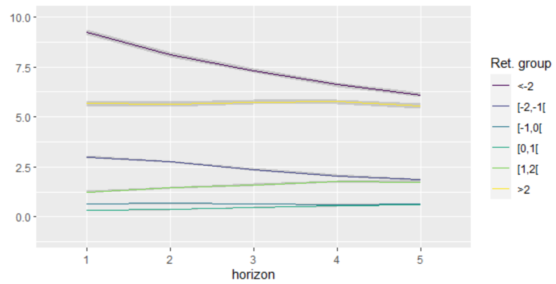
Panels A and C display the average RH users change with 95% confidence bands as a function of group of returns level or horizon, as described in section 4.1 ($\mu_{g,h}$). Panels B and D display results from the estimation of regression (4), as described in section 4.2 ($\beta_{g,h}$). The RH users change variable, $\Delta N_{i,t_{i,k}}$ is winsorized at the 95% and 5%. All estimates are expressed in basis point.



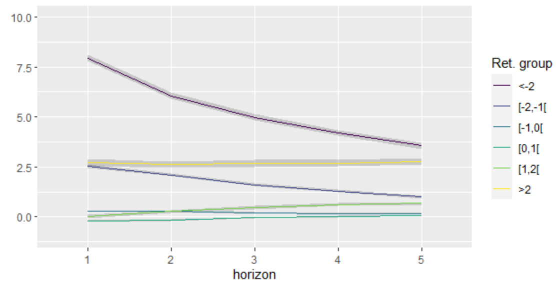
(a) $\mu_{g,h}$ by group of return levels



(b) $\beta_{g,h}$ by group of return levels



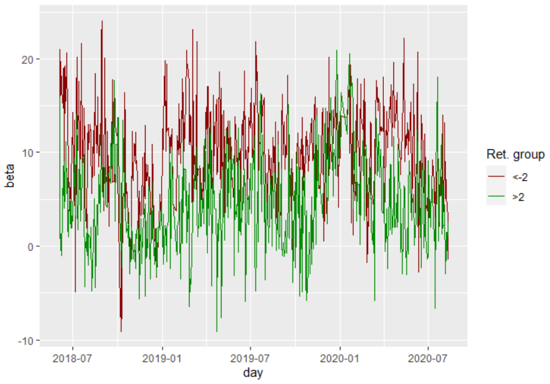
(c) $\mu_{g,h}$ by horizon



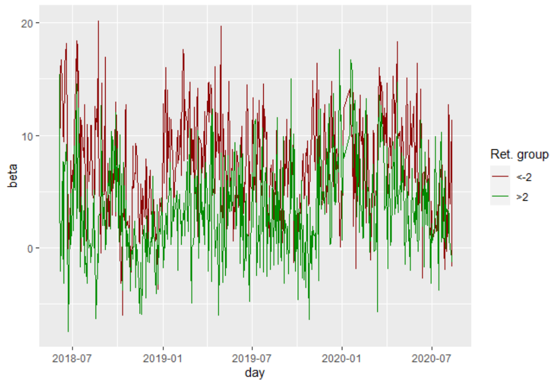
(d) $\beta_{g,h}$ by horizon

Figure 2: Time-series variations of RH investors' reaction to past extreme returns

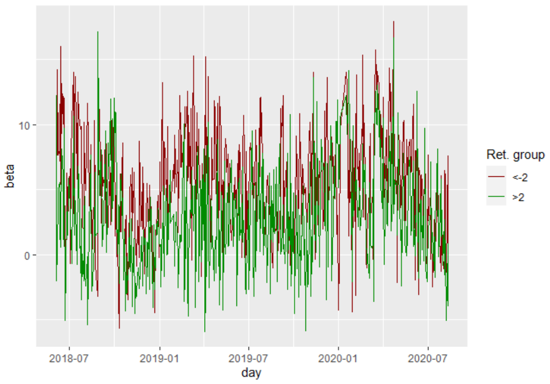
This figure displays the $\hat{\beta}^{(d)}_{g,h}$ from regression (4) run separately for each day corresponding to groups of extreme return levels (*i.e.*, $\mathcal{G}_1$ and $\mathcal{G}_6$), at one- to five-hour horizon. Estimates are expressed in basis point.



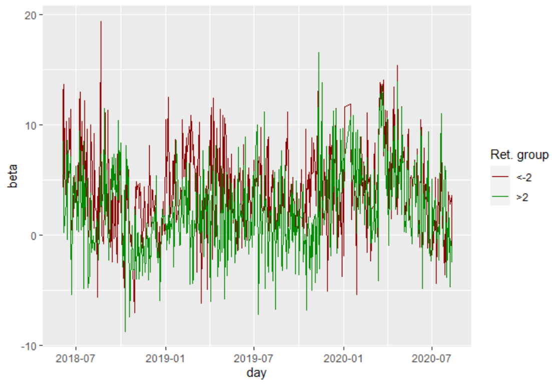
(a) One-hour horizon



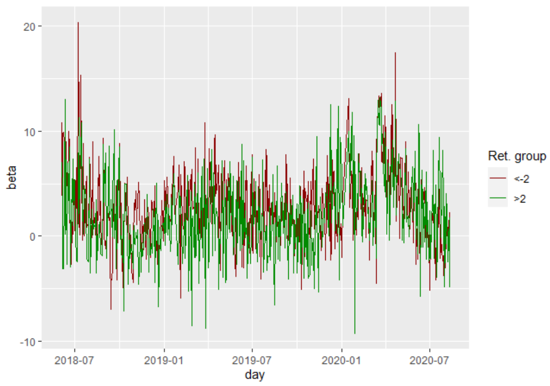
(b) Two-hour horizon



(c) Three-hour horizon



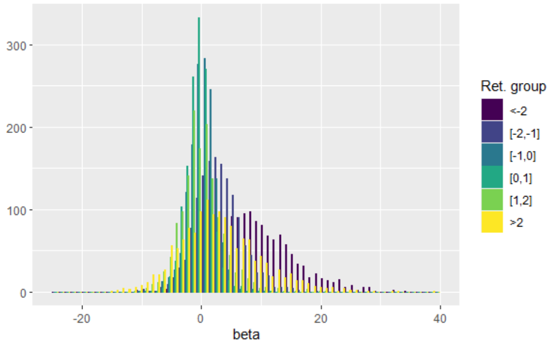
(d) Four-hour horizon



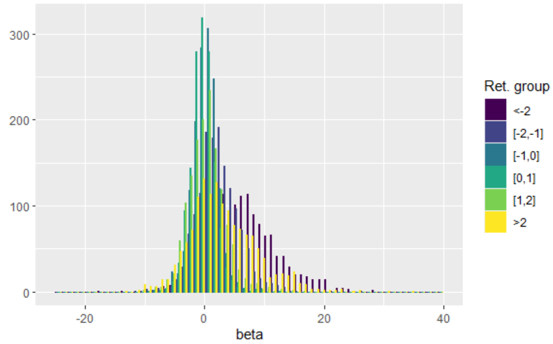
(e) Five-hour horizon

Figure 3: Cross-sectional variations in RH investors' reaction to past returns

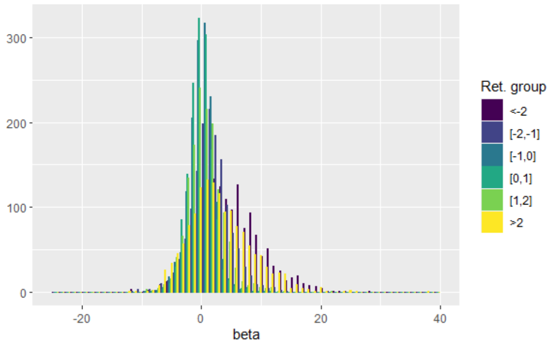
This figure displays the distribution of $\hat{\beta}_{g,h}^{(i)}$ based on regression (4) run separately for each company corresponding to all groups of return levels ($i.e.$, $\mathcal{G}_1, \ldots, \mathcal{G}_6$), at one- to five-hour horizon. Estimates are expressed in basis point.



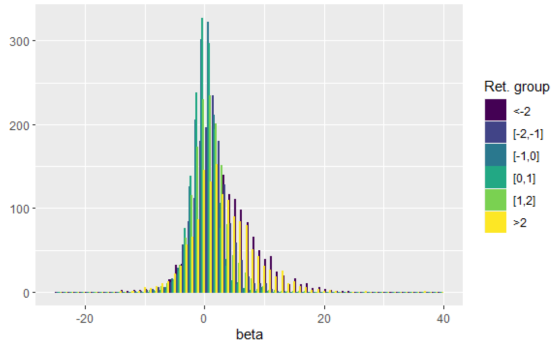
(a) One-hour horizon



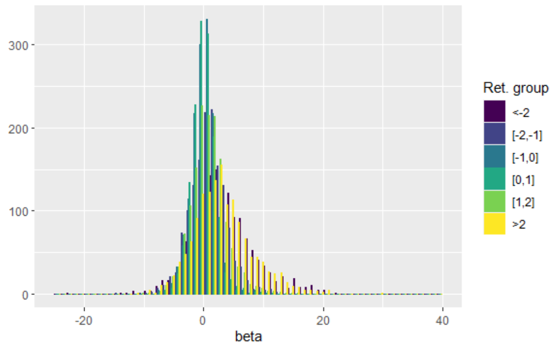
(b) Two-hour horizon



(c) Three-hour horizon

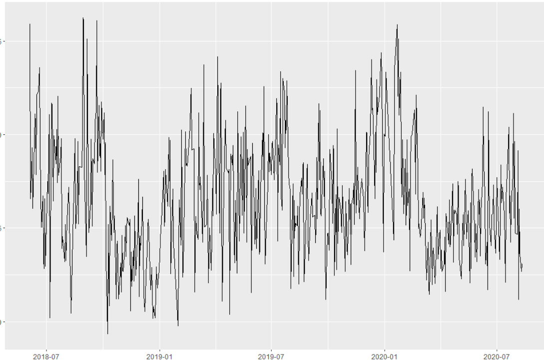


(d) Four-hour horizon

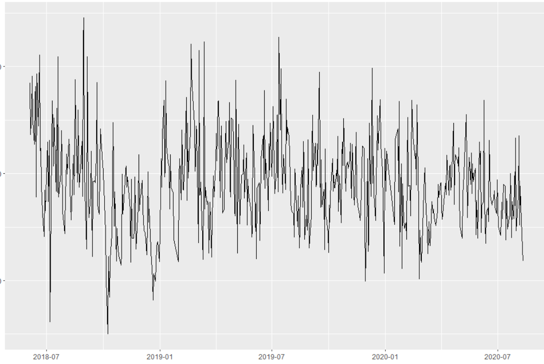


(e) Five-hour horizon

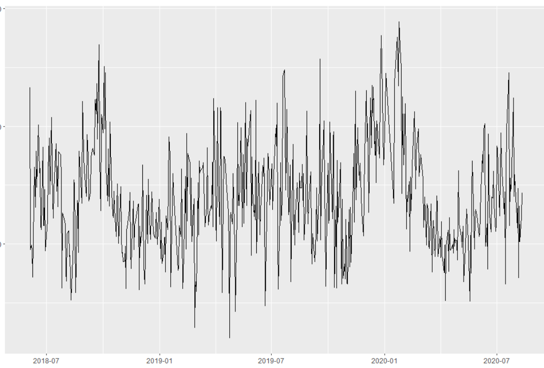
Figure 4: Proxy variables representing RH investors' behavioral patterns in the time series
This figure presents the daily series of our proxy variables representing key RH investors' behavioral patterns with respect to past returns. Proxies from Panels A, B, C, D and E are defined in (5), (6), (7), (8) and (9), respectively.



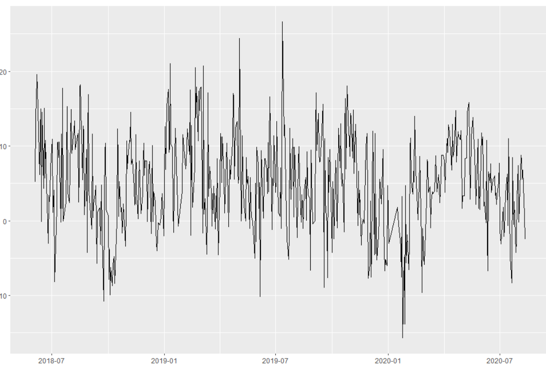
(a) $Ext^{(d)}$



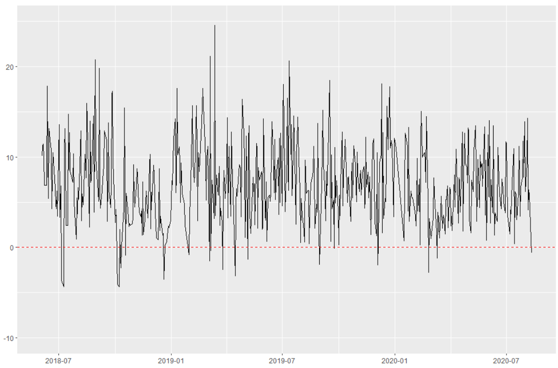
(b) $ExtNeg^{(d)}$



(c) $ExtPos^{(d)}$



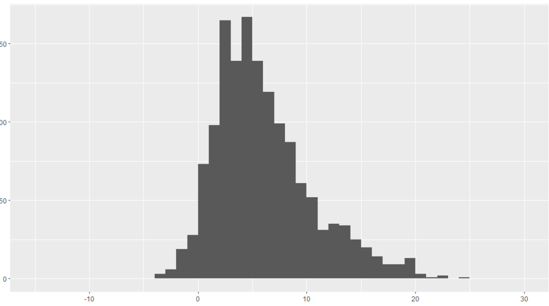
(d) $Asy^{(d)}$



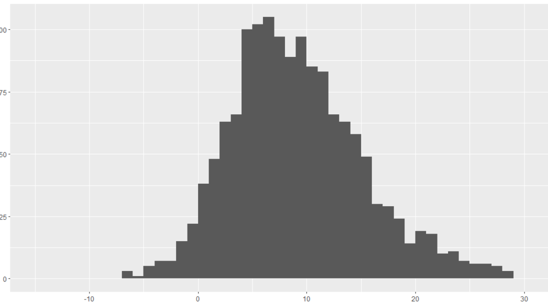
(e) $SpeedExtNeg^{(d)}$

Figure 5: Proxy variables representing RH investors' behavioral patterns across firms

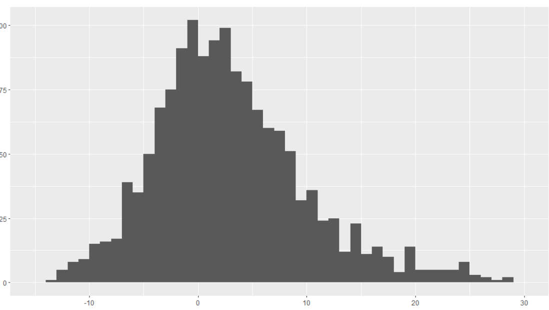
This figure presents the distribution across companies of our proxy variables representing key RH investors' behavioral patterns with respect to past returns. Proxies from Panels A, B, C, D and E are defined analogously to equations (5), (6), (7), (8) and (9), respectively.



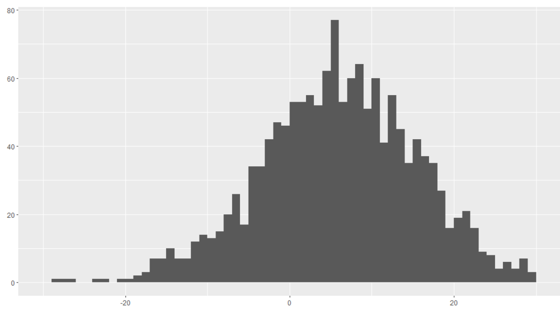
(a) $Ext^{(i)}$



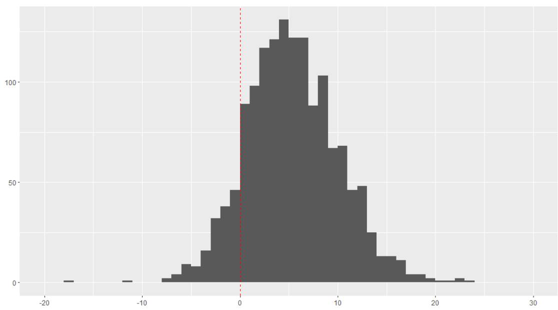
(b) $ExtNeg^{(i)}$



(c) $ExtPos^{(i)}$

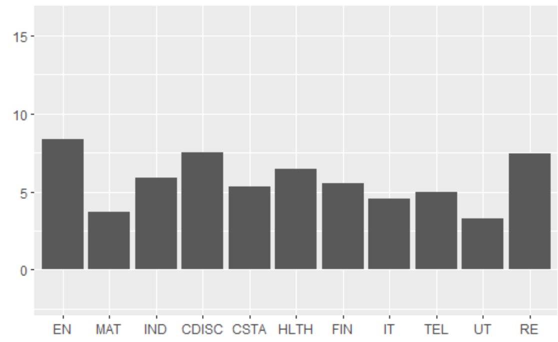


(d) $Asy^{(i)}$

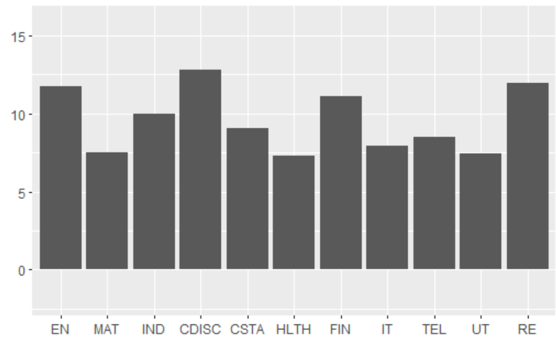


(e) $SpeedExtNeg^{(i)}$

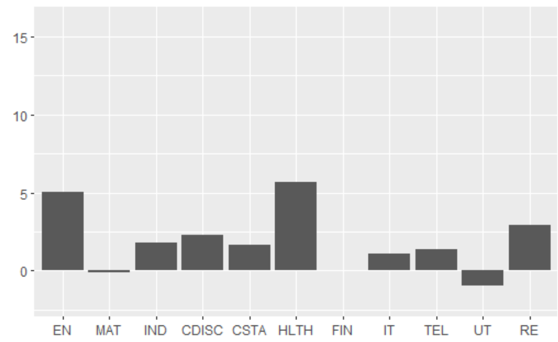
Figure 6: RH investors' behavior with respect to previous returns per industry
This figure displays the average of the proxy variable defined in section 6.2 conditional on the firm being in a given sector, expressed in basis point. The averages are estimated within a regression setup. For more detail on the estimation results, refer to Section A.5 of the appendix.



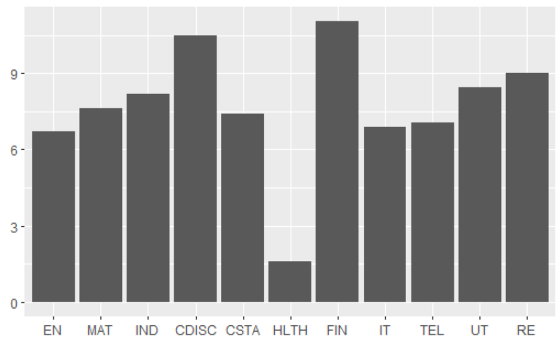
(a) $Ext^{(i)}$



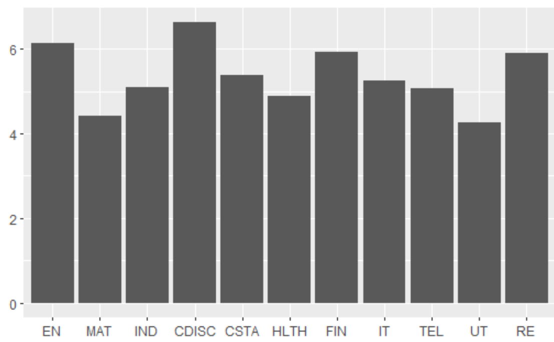
(b) $ExtNeg^{(i)}$



(c) $ExtPos^{(i)}$



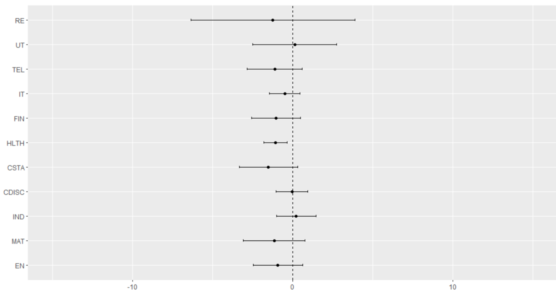
(d) $Asy^{(i)}$



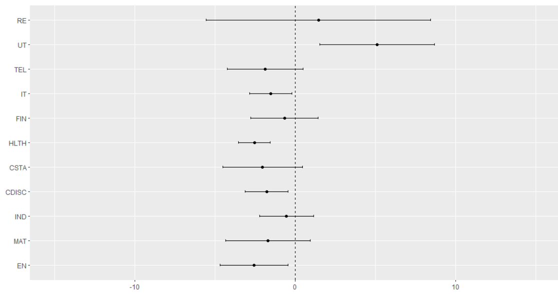
(e) $SpeedExtNeg^{(i)}$

Figure 7: Effect on COVID-19 on the RH investors' behavior per industry

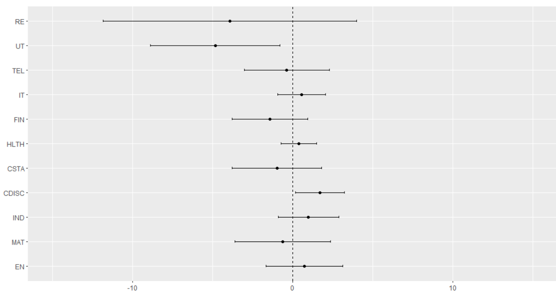
This figure displays the differences between the average of each proxy variable conditional on the firm being in a given sector taken after and before the COVID-19 pandemic announcement. Proxy variables are defined in section 6.2. All averages are estimated within a regression setup. For more details on the estimation results, refer to Section A.6 of the appendix. The points indicate the difference between the averages after and before the COVID-19 pandemic announcement. The bars represent the 95% confidence intervals around these differences resulting from F-tests on coefficient equality (*i.e.*, for each sector S, $H_0 : \hat{\beta}_{S.preCOV} = \hat{\beta}_{S.postCOV}$). All values are expressed in basis point.



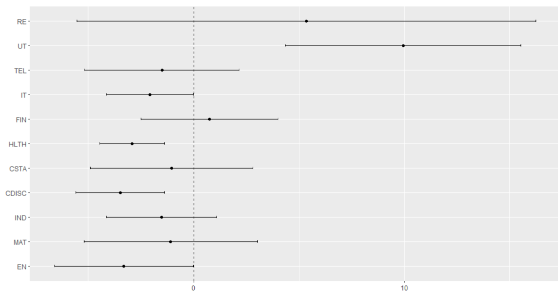
(a) $Ext^{(i)}$



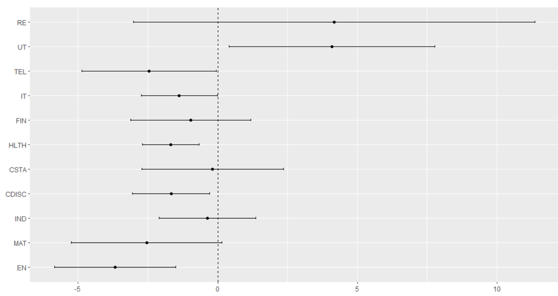
(b) $ExtNeg^{(i)}$



(c) $ExtPos^{(i)}$



(d) $Asy^{(i)}$



(e) $SpeedExtNeg^{(i)}$