How do Robinhood investors trade? An intraday-level study*

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*

Very Preliminary

Abstract

We propose a comprehensive study on the behavior of Robinhood investors with respect to past returns at the intraday level. We find 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 with sharp declines, and, (iii) Robinhood users tend to respond particularly fast —after only one hour— to previous large negative returns. These results can be partially explained by the level of market volatility or investors sentiment as well as firm characteristics such as financial health, liquidity, or specific risk.

Keywords: Retail investors, investor attention, Robinhood, behavioral finance, intraday

---

*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)

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 picked to more than 40% during the first quarter of 2021, from only 25% in mid-2018 (see The Economist, 2021). The online trading platform offering zero trading commissions, "Robinhood Market Inc.", founded in 2013, 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 18 million in 2021, almost three times its level of 2018 (see BusinessofApps, 2021).

Robinhood investors are now able to move the market. Examples are multiplying. On May 22, 2020, the rental car company Hertz filed for bankruptcy. On May 26, its stock price felt 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 a tiny portfolio, with an estimated average account size between \$3,000 and \$5,000, and they trade mainly through the RH app, which has often been criticized for its design prone to the "gamification" of investing.[1]

In this paper, we examine the trading behavior of this new type of investor. A key element of our research is that we investigate RH investors' behavior at the *intraday* level. In contrast, the closest papers to ours, Barber et al. (2021) and Welch (2021), provide results at the daily frequency that we believe tell only part of the story. Indeed, these so-called Generation Z investors who grew up in the internet and social media era are ultra-connected. Within a single day, they are exposed to considerable information. Besides, as supported by Barber et al. (2021), they are highly influenced by the app's design, which can send notifications anytime. Therefore, they likely adjust their positions more than once a day. Studying their behavior on a finer scale appears highly relevant in that context. To the best of our knowledge, we are the first to provide results at this frequency level.

For simplicity, and building on Barber and Odean (2008), we choose previous (hourly) returns as a proxy for RH investors' attention. While we acknowledge that this proxy for

[1]This typical RH investor profile has been drawn multiple times in press articles (*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 (*e.g.*, Barber et al., 2021; Welch, 2021; Eaton et al., 2021; Van der Beck and Jaunin, 2021).

attention is imperfect, we believe it is a good candidate given the inexperienced nature of RH investors. An investor with no experience is likely to be influenced by the most straightforward and visible information he is subject to: extreme returns. While other studies are devoted to finding the best proxy for the attention of individual investors[2], this is not the focus of this paper. We instead assume that this proxy for attention is relevant and use it to develop and test the following hypotheses empirically:

H1 *RH investors react stronger to stocks exhibiting past extreme returns.* Barber and Odean (2008), using private data from brokerage firms, show that this statement is valid for the entire population of retail investors and on a daily basis. With regards to RH investors, in particular, Barber et al. (2021) find that extreme daily returns can be one of the factors explaining herding events. Welch (2021) find that "RH investors in 2020 liked to purchase both large gainers and large losers", using extreme daily returns as well. We go one step further by assuming that this phenomenon is also true at a higher (hourly) frequency level.

H2 *The reaction to previous extreme returns is not symmetrical.* We posit that the strength of the reaction is different depending on the direction (sign) of the extreme return.[3] This question can relate to the style of RH investors, *e.g.*, are they more contrarian investors (buy [sell] more stocks exhibiting large negative [positive] returns), momentum investors (buy [sell] more stocks exhibiting large positive [negative] returns), or both?[4]

H3 *RH investors' speed of reaction to past extreme returns is high.* Based on the rationale developed above (ultra-connectedness, information sharing through Reddit chats, intrusive app), we test the speed of reaction of RH investors. We assume that they should be made aware of the event (in our case, the previous extreme returns) quite early, so their reaction must be stronger the first hours following the event. To test this claim, we compare the strength of the responsiveness to an extreme return after one to five hours.

Using a comprehensive dataset from Robintrack.net and intraday prices from the NYSE Trade And Quote (TAQ) millisecond datasets, we test these hypotheses empirically. Our results confirm at least partially all three hypotheses. For *H1*, we find that, on average, the reaction to large past hourly returns is stronger than the reaction to moderate past hourly returns. Therefore, the results of Barber and Odean (2008), which covers all retail

[2]Among others, Kaniel and Parham (2017) use news from the Wall Stree Journal, Da et al. (2011) and Andrei and Hasler (2015) advocate for Google search data, and Cookson and Niessner (2020) exploit messages from the StockTwits platform. While papers studying individuals' trading behavior have identified a link with IQ level (Grinblatt et al., 2011), social influence (Shive, 2010), or herding effect (Kumar and Lee, 2006).

[3]Barber and Odean (2008) find mixed results: asymmetry is more or less pronounced depending on the proxy (brokerage firms) used.

[4]One can also relate this question to some behavioral biases identified in the literature, in particular the disposition effect (Shefrin and Statman, 1985), which is the tendency to sell past winners and hold on past losers.

investors and daily returns, also hold for a particular type of individual investors (the RH investors) and at the intraday level. With regards to *H2*, we find that, on average, the reaction is asymmetrical to the favor of large previous negative returns, *i.e.*, the "net buying" pattern is stronger for stocks exhibiting large negative returns than for stocks exhibiting large positive returns. Finally, for *H3*, we observe that, on average, the speed of reaction to past extreme *negative* returns is high. More precisely, we find that the "net buying" pattern is more pronounced when we evaluate the reaction one hour following a large negative return compared to the reaction two, three, four, or five hours following the event. However, that it is not necessarily the case for past extreme positive returns. Therefore, *H3* is only partially confirmed. This result suggests that RH investors are (i) made aware of the occurrence of a negative stock return relatively fast and (ii) tend to act on this information quite fast as well.

The second part of this paper builds on the empirical results described above. These results, which are based on averages across all stocks and all days, are revisited in time series and cross section analyses. The goal is to identify how these patterns look in both dimensions, and to explore the determining factors that might guide them. We first construct proxies representing the behavior of RH investors derived from our hypothesis tests. Then, we use each proxy as a dependent variable and regress it on time series factors and firm characteristics. For time series variations, we assess the potential impact of market volatility, excess market returns, and investor sentiment. Our main findings are that market volatility plays an important role in explaining the different patterns of RH investors' behavior over time. We also find that investor sentiment is strongly correlated with our RH investors' behavior proxies. To explain cross-sectional variations, we regress our proxies on various firm characteristics. For some RH investors' behavioral patterns, such as their reaction to extreme returns, variations across firms can be partially explained by their financial health, trading volume, or liquidity level. For other RH investors' behavioral patterns, such as their speed of reaction to extreme negative returns, variations can be partially explained by firm size, systematic risk, or idiosyncratic risk.

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 does give some insights about the behavior of RH investors 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 (time period where RH investors can not trade due to technical breakdowns) the overall activity of retail investors[5] is significantly reduced. This shows evidence that RH investors represent a large chunk of total individual investors activity, and reinforces 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 has impacts on 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 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) show that RH investors did not particularly care about ESG investing. Ozik et al. (2021) show that during the COVID-19 pandemic 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

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

pandemic locking on RH investors but through a different lense: 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 stuctural model to quantify the impact of Robinhood traders on the US equity market.

The rest of the paper is organized as follows. Section 2 discusses our data sources and define our main variables. Section 3 presents the results on the empirical tests of our three hypotheses. Section 4 investigate these results both in the time series and cross section, and explore the potential factors explaining them. Section 5 concludes.

2. Data and variable definitions

2.1. Robintrack

From mid–2018 to mid–2020, it was possible to use Robinhood's API to obtain data 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 (Fortune, 2020). This paper uses RT's data. A clear definition of the data and its limitations is important. On its website, RT describes it as *"how many Robinhood users hold a particular stock over time"*, and advises to *"keep in mind that this data is limited in a few ways: It only shows when users buy shares after holding none rather than when they add shares to existing positions [...]"* and adds that *"Also, it is heavily influenced by the number of new Robinhood accounts by either making their first trades or receiving free shares as part of the referral program"*. Hence, each observation does not represent the *number of shares* held by RH investors, but the *number of RH investors who hold at least one share* of a given stock.[6]

The original database, retrieved directly from RT's website, is very rich. It spans the period from May 1, 2018 to August 13, 2020, and contains more than 140 million observations at the hourly frequency, on more than 8,000 distinct securities (tickers). Because these data are not from a professional source, we apply some adjustments. In particular, we follow Welch (2021) and drop of the first month of the original period, May 2018, as we suspect the reliability of the data improves as time passes. Since Robinhood publishes the number of users holding a stock every hour with an approximate 45-minutes delay, we subtract 45 minutes from every observations' timestamps.[7] To match RH observations with prices data,

[6]For example, if the number of RH users for Apple Inc. (AAPL) increases between two observations (as represented by timestamps), 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. In other words, we cannot state that the total number of shares of AAPL held by Robinhood traders is higher, as it is possible that during this time interval, a group of RH investors reduced or added from an already existing position in the stock.

[7]Our discussion with Casey Primovic, the administrator of Robintrack, led us to estimate the timestamp

we also reduce the number of observations to be within market-opening hours (9.30am to 4.00pm) only. We reduce further the number of securities to common stocks only (CRSP share codes of 10 or 11). The detailed list of adjustments and their impact on the original data-set can be consulted in the appendix.

2.2. Other data

We also use transaction prices from the NYSE Trade And Quote (TAQ) millisecond datasets. For each stock, we extract (i) its last trade price available before the timestamp corresponding to the RT observation and (ii) the last price of the SPDR S&P 500 ETF (SPY), which is our market proxy. To ensure reliability of these prices, we use the filter rules described in Barndorff-Nielsen et al. (2009)[8].

The second set of results uses additional data. For the time-series analyses, we use the Cboe Volatility Index (VIX) series from CRSP, the daily excess market returns series from Kenneth R. French's website, and the sentiment index series from the American Association of Individual Investors. For the cross-section analyses, we construct various firm characteristics using accounting and industry data from COMPUSTAT, and stock related data from CRSP [9].

2.3. Variable definitions

First, we build our proxy for RH investors' trading behavior. This variable, ΔRH, measures the change in the number of RH investors holding a given stock at the hourly frequency. Because we keep only observations within market-opening hours, we must differentiate the change between two intraday observations (*e.g.* 2018-06-04 09:45 and 2018-06-04 10:44) from the (overnight) change between two consecutive trading days observations (*e.g.* 2018-06-04 15:45 and 2018-06-05 09:43). Therefore, we create two separate measures that are consistent with each other. An intraday change is measured as follows:

$$\Delta RH_{i,t,\widehat{h}} = \log\left(\frac{Nusers_{i,t,h}}{Nusers_{i,t,h-1}}\right)\frac{60}{Nmin(h,h-1)}, \tag{1}$$

where *Nusers* is the number of RH investors detaining stock i at time h of day t.[10] Because two consecutive observations are not necessarily spaced by exactly 60 minutes, the second term is a one-hour scaling factor where $Nmin(h, h-1)$ is the number of minutes between time $h-1$ and h. As emphasized by the subscripts t in the first term, this measure is computed

lag between 30 and 60 minutes. Therefore, a data point with an associated timestamp of 10:45 actually represents a snapshot of the actual data point 30 to 60 minutes before that time. This paper assumes a timestamp lag of 45 minutes. For robustness check, we also provide results using timestamp lags of 30 and 60 minutes (see appendix, XXX). A discussion of this issue is also provided in Barber et al. (2021).

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

[9] data fields used from COMPUSTAT: *MKVALTQ*, *CEQQ*, *DLCQ*, *DLTTQ*, *ATQ*, *GSECTOR*. Data fields used from CRSP: *vol*, *bid*, *ask*. Firm characteristics are defined in section 4.

[10] to avoid 0 in the denominator of $\log\left(\frac{Nusers_{i,t,h}}{Nusers_{i,t,h-1}}\right)$, we added 1 to the original *Nusers* measure.

only for two consecutive *same-day observations*. For an overnight change, we define the following measure:

$$\Delta RH_{i,t,\tilde{h}} = \log\left(\frac{Nusers_{i,t,firsth}}{Nusers_{i,t-1,lasth}}\right)\frac{60}{Nmin(firsth, lasth)}\,, \tag{2}$$

where $Nusers_{i,t,firsth}$ ($Nusers_{i,t-1,lasth}$) is now the number of investors detaining stock i at the $first$ ($last$) timestamps available on day t ($t-1$). The second term scales the measure to one-hour change where $Nmin(firsth, lasth)$ is the number of minutes between the last observation of day $t-1$ and the first observation of day t. As emphasized by the subscripts t and $t-1$ in the first term, this measure is computed only for two consecutive *different-day observations*.

We obtain our measure of change in RH investors' behavior by combining (1) and (2):[11]

$$\Delta RH_{i,t,h} = \begin{cases} \Delta RH_{i,t,\widehat{h}} & \text{for an intraday change} \\ \Delta RH_{i,t,\tilde{h}} & \text{for an overnight change} \end{cases}. \tag{3}$$

We now turn to our high-frequency returns proxy measure. To better capture the true extreme movements, we use an adjusted measure by scaling each stock (log) returns by their respective volatility over the entire period. As above, we differentiate intraday returns from overnight returns. For the computation of adjusted intraday returns, let's first define intraday *unadjusted* log-returns scaled to one-hour and their corresponding volatility as:

$$r_{i,t,\widehat{h}} = \log\left(\frac{p_{i,t,h}}{p_{i,t,(h-1)}}\right)\frac{60}{Nmin(h, h-1)} \tag{4}$$

$$\sigma_i(r_{i,t,\widehat{h}}) = \sqrt{\frac{\sum_{\widehat{h}=1}^{H(t)}\sum_{t=1}^{T(i)}\left(r_{i,t,\widehat{h}} - \bar{r}_i\right)^2}{\sum_{t=1}^{T(i)} H(t) - 1}}\,, \tag{5}$$

where $p_{i,t,h}$ is the price of stock i at time h of day t, $H(t)$ is the number of observations in day t, and $T(i)$ is the number of days for which stock i has observations. The *adjusted* one-hour scaled intraday returns measure is then:

$$\tilde{r}_{i,t,\widehat{h}} = \frac{r_{i,t,\widehat{h}}}{\sigma_i(r_{i,t,\widehat{h}})}\,. \tag{6}$$

We compute our measure of adjusted overnight (log) returns in the spirit of Lou et al. (2019). The unadjusted overnight returns and their corresponding volatility over the entire

[11]Figure XXX in the appendix shows that these two measures are directly comparable to each other.

period are respectively[12]:

$$r_{i,t,\tilde{h}} = \log\left(\frac{p_{i,t,lasth}}{p_{i,t-1,lasth}}\right) - \log\left(\frac{p_{i,t,lasth}}{p_{i,t,firsth}}\right) = r_{i,t}^{CLOSE-TO-CLOSE} - r_{i,t}^{CLOSE-TO-OPEN} \tag{7}$$

$$\sigma_i(r_{i,t,\tilde{h}}) = \sqrt{\frac{\sum_{t=1}^{T(i)}\left(r_{i,t,\tilde{h}} - \bar{r}_i\right)^2}{T(i)-1}}\,. \tag{8}$$

The adjusted overnight returns measure is then:

$$\tilde{r}_{i,t,\tilde{h}} = \frac{r_{i,t,\tilde{h}}}{\sigma_i(r_{i,t,\tilde{h}})}\,. \tag{9}$$

We obtain our measure of adjusted returns by combining (6) and (9):[13]

$$\tilde{r}_{i,t,h} = \begin{cases} \tilde{r}_{i,t,\hat{h}} & \text{for an intraday return} \\ \tilde{r}_{i,t,\tilde{h}} & \text{for an overnight return} \end{cases}\,. \tag{10}$$

Our final sample consists of 8,03 million observations on 2853 stocks and 527 trading days from June 1st, 2018 to August 13, 2020. Panels A and B of Table 1 present summary statistics on our RH trading behavior ($\Delta RH_{i,t,h}$) and adjusted return measures ($\tilde{r}_{i,t,h}$). The statistics are taken across all stock-hour observations. The average (one-hour scaled) RH users change is positive and rather high, at 4.36 basis points. This is due to the success of Robinhood: during the period, the number of users was almost constantly increasing. 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 our hourly observations do not change from one hour to another, the median change is null. The distribution of our measure appears to be right-skewed, in particular due to extreme positive values. The average one-hour scaled RH users change overnight and intraday are at 4.09 and 4.41 basis point respectively. Based on the standard deviation and quartiles, overnight changes seem to be less volatile and more right-skewed (for an histogram representation of distributions, refer to Appendix, Figure XXX).

Not surprisingly, Panel B shows that our adjusted returns measures have approximately 0 mean and standard deviations 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.

[12] For the purpose of this study, we substitute the close (open) price by the price corresponding to the last (first) RH observation before 4.00pm (after 9.30am)

[13] Figure XXX of the appendix shows that the distributions of adjusted intraday log-returns and adjusted overnight log-return are similar.

3. Empirical results

Using the two measures defined above, we now turn to the empirical tests of the three hypotheses developed in the introduction, namely:

H1: RH investors react stronger to stocks exhibiting past extreme returns.

H2: The reaction to previous extreme returns is not symmetrical.

H3: RH investors' speed of reaction to past extreme returns is high.

3.1. Preliminary evidences

We start with a simple approach. First, we classify stock adjusted returns into group levels. Then, we compute the average subsequent change in RH users evaluated at different horizons for each group level. More formally, let's define the following indicator function:

$$I(g) = \begin{cases} 1, & \text{if the return belongs to group } g \\ 0, & \text{otherwise} \end{cases}$$

where $g \equiv \{< 2, [-2, -1[, [-1, 0[, [1, 2[, > 2\}$ is the set of return group levels. The indicator we are interested in is:

$$E\left[\Delta RH_{i,t,h} | \tilde{r}_{i,t,(h-x)}.I(g)\right] , \tag{11}$$

which designs the average RH users change at the x hour(s) horizon (*i.e.*, the number of hours passed between the return and the change in RH users), conditional on the previous adjusted returns being in a certain group level g. We compute this measure over all stock-hour observations for the six return group levels, and at the one to five hours horizon (*i.e.*, $x = 1, 2, \ldots, 5$), which result in a total of 30 conditional averages. Figure 1 displays the results. Panel A, which formats the results using return group levels (g) as the x-axis, seems to support *H1* and *H2*. Indeed, we first observe that the average change in RH users is U-shaped: RH investors tend to open more new positions in stocks which just had large stock price movements, whether positive or negative, than in stock exhibiting previous moderate price movements. For example, a large negative adjusted return (g = <-2) has produced an average RH users change of 9.23 basis point during first the hour ($h = 1$) following this return. This is significantly higher than the average RH users change of 0.33BP subsequent to a moderate adjusted return ($g = [0, 1]$). Similarly, a large positive adjusted return (g = >2) has produced an average RH users change of 5.66BP during the second hour ($h = 2$) following this return. This is significantly higher than the average RH users change of 0.37BP subsequent to a moderate adjusted return ($g = [0, 1]$).

Second, as assumed in *H2*, the reaction appears asymmetric. We actually observe a smirk: the average change in RH users following a large *negative* return is superior to the average change in RH users following a large *positive* return. For example, at $h = 1$, the average RH users change conditional on past large negative return (9.23BP) is more than 60% superior to its counterpart conditional on past large positive return (5.67BP). We also note that *H1* and *H2* seem to hold at all horizons. To help interpret results relative to *H3*, Panel B uses the horizon as its x-axis. The average reaction to large *negative* adjusted returns appears to decrease monotonically as the horizon increases. The average reaction to extreme *positive* returns, however, do not seem to change with the horizon. Therefore, it suggests that *H3* only holds for the negative side of the extreme returns. Interestingly, the means are also monotonically decreasing for the $[-2,-1[$ returns group, but not for any other groups. This suggests that RH investors are particularly quick at taking action in stocks exhibiting past, and not necessarily extreme, negative returns.

3.2. 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 x hour(s) after a given adjusted return, one might want to verify if it is not caused by another adjusted return occurring before or after this return. For example, let's say we are interested in the reaction of RH investors $x = 1$ hour after a given return. In that case, we will control for the returns preceding the reaction of RH investors at two to five hours. Similarly, if we assess the reaction of RH investors $x = 2$ hours after a given return, we will control for the reaction of RH investors at one and three to five hours. Another potentially important factor is the market return as the reaction might be market-wide and not stock specific. We therefore control for market returns as well, lagged by one to five hours relative to the time of the reaction of RH investors. We also add a quadratic version of these controls to account for non-linearity. We specify the regression as follows:

$$\begin{aligned} \Delta RH_{i,t,h} = & \sum_{g=1}^{G} \beta_g \tilde{r}_{i,t}(k=x).I(g) + \sum_{k \in K} \gamma_k \tilde{r}_{i,t}(k) + \sum_{k \in K} \delta_k \tilde{r}_{i,t}(k)^2 \\ & + \sum_{k \in K_m} \psi_k \tilde{r}_{m,t}(k) + \sum_{k \in K_m} \xi_k \tilde{r}_{m,t}(k)^2 + \epsilon_{i,t,h} \, , \end{aligned} \tag{12}$$

where g designs the return level group, $I(g)$ is the indicator variable as defined above and G is the number of groups. $\tilde{r}_{i,t}(k)$ is the stock i adjusted return of day t and k represents the lag relative to the time at which the RH investor reaction is evaluated (h).[14] All other terms relate to our control variables. The set of terms from the second and third summation control for stock i other adjusted returns with different lags up to five hours, as represented by $K = \{1 \leq k \leq 5, k \neq x\}$. The set of terms from the fourth and fifth summation control for market adjusted returns $(\tilde{r}_{m,t}(k))$ lagged by one to five hours, as represented by $K_m = \{1 \leq k \leq 5\}$.

[14]For example, $\tilde{r}_{i,t}(k=1)$ is the return one hour before h, therefore such a specification would evaluate the reaction of RH investors at the one hour horizon. Similarly, $\tilde{r}_{i,t}(k=2)$ is the return two hours before h, therefore such a specification would evaluate the reaction of RH investors at the two hours horizon, and so on.

We are interested in the β_g coefficients which, as in the previous section, represent the mean reaction of RH investors to stock adjusted returns from group level g at a given horizon x. However, this mean is now freed from the effect of the control variables. We estimate this regression for each of our five horizons, when $x = 1, ..., 5$. Table 2 displays estimation results. All five regressions are based on the full panel and are estimated by standard OLS. The t-statistics are computed with Newey and West (1987) standard errors estimators. Adding the control variables do not significantly impact the results suggested by the conditional means approach. Out of the 30 β_g estimates, 28 are significant at the 1% level. More important, all the β_g coefficients evaluating RH investors reaction to *extreme* returns are highly significant. As Figure 2 shows, like in the previous section, the U-shape pattern is still present, giving further support for *H1*. Moreover, the asymmetric characteristic of the smile is even more pronounced when we include the control variables, which strengthens the validity of *H2*. Comments made in the previous section relative to *H3* are still valid, but compared to the conditional mean approach, the monotonically decrease in the average reaction to large negative returns as a function of horizon is now even stronger. Therefore, RH investors' speed of reaction to large negative returns seems to be higher after controlling for other stock returns and market returns.

3.3. Are these patterns consistent in the time-series and cross-section?

We first look at these results on a daily basis over our entire sample period, from June 1st, 2018 to August 13, 2020. We run Regression (12) separately for each day:

$$\begin{aligned} \Delta RH_{i,h}^{(t)} = &\sum_{g=1}^{G} \beta_g^{(t)} \tilde{r}_{i,t}(k=x).I(g) + \sum_{k \in K} \gamma_k^{(t)} \tilde{r}_{i,t}(k) + \sum_{k \in K} \delta_k^{(t)} \tilde{r}_{i,t}(k)^2 \\ &+ \sum_{k \in K_m} \psi_k^{(t)} \tilde{r}_{m,t}(k) + \sum_{k \in K_m} \xi_k^{(t)} \tilde{r}_{m,t}(k)^2 + \epsilon_{i,h}^{(t)} . \end{aligned} \tag{13}$$

With 526 trading days and five horizons, we estimate a total of 2630 regressions. On average, each daily regressions have 15,239 observations. Table XXX in appendix summarizes the results and Figure 3 plots the time series of the estimated daily betas corresponding to group < -2 and > 2. We can observe that the coefficients fluctuate over time. Therefore, the results described above – that (*H1*) RH investors respond stronger to extreme returns than to moderate returns, (*H2*) RH investors respond stronger to large negative returns than to large positive returns, and (*H3*) RH investors' speed of reaction is higher following large negative returns – must also vary over time.

Similarly, we investigate how these results look in the cross-section. Is there lot of heterogeneity across companies? For a reliable analysis, we first ensure that the dependent variable, $\Delta RH_{i,t,h}$ has enough observations (at least one year) for each stock i. We also require that it exhibits sufficient fluctuations, *i.e.*, that the company is an "active Robinhood stocks" in the sense that the number of RH users change is different than zero for more than 50% of observations. After these adjustments, our sample reduces to 1,454 companies. For each company, we run regression (12) for each horizon $x = 1, \ldots, 5$,

$$\begin{aligned}\Delta RH_{t,h}^{(i)} = \sum_{g=1}^{G} \beta_g^{(i)} \tilde{r}_{i,t}(k=x).I(g) + \sum_{k \in K} \gamma_k^{(i)} \tilde{r}_{i,t}(k) + \sum_{k \in K} \delta_k^{(i)} \tilde{r}_{i,t}(k)^2 \\ + \sum_{k \in K_m} \psi_k^{(i)} \tilde{r}_{m,t}(k) + \sum_{k \in K_m} \xi_k^{(i)} \tilde{r}_{m,t}(k)^2 + \epsilon_{t,h}^{(i)}\end{aligned} \tag{14}$$

leading to a total number of 7,270 regressions. On average, each company regression has 3,089 observations. Table XXX in the appendix summarizes the results and Figure 4 displays the distribution of the cross-sectional regression betas at each horizon. The distributions have strong dispersion. For example, the company $\beta_{<-2}$ have an average value of 8.95BP and a standard deviation of 6.54BP at the one-hour horizon. For the reaction to other return group levels, the dispersion is even stronger (*e.g.*, at the one-hour horizon, $\bar{\beta}_{>2} = 2.73$BP and $\sigma(\beta_{>2}) = 6.86$BP). Therefore, there appears to be heterogeneity across firms in the results.

The variation of the results in the time-series and cross-section motivate the analyses performed in the next section.

4. In search of the determinants of RH investors' behavior

Since the results vary across time and companies, the natural question is whether it is possible to explain these variations. This section is devoted to studying the potential factors driving these variations. We proceed as follows. First, based on the estimated coefficients from the time-series and cross-sectional regressions presented in section 3.3, we construct proxy variables representing the RH investors' behavioral patterns identified in our hypothesis test. Then, we regress these proxies on time-series factors or firm characteristics. For example, with regards to *H1*, we have found evidence that on average, RH investors respond stronger to past extreme returns but that these responses can vary through time and among companies. In a first step, we define proxy variables representing RH investors reaction to extreme return. In a second step, we run time series and cross-sectional regressions with these proxies as dependent variables and the time series factors or firm characteristics as independent variables. We proceed similarly for results associated to *H2* and *H3*.

4.1. Setup

4.1.1. Proxy variables

To evaluate *H1* and *H2*, we construct the following variables:

$$Y_1^h = \frac{\beta^h_{g=<\text{-}2} + \beta^h_{g=>2}}{2} - \frac{\beta^h_{g=[\text{-}1,0]} + \beta^h_{g=[0,1]}}{2} = \bar{\beta}_{\text{Ext}} - \bar{\beta}_{\text{Mod}} \tag{15}$$

$$Y_2^h = \beta^h_{g=<\text{-}2} - \frac{\beta^h_{g=[\text{-}1,0]} + \beta^h_{g=[0,1]}}{2} = \bar{\beta}_{\text{ExtNeg}} - \bar{\beta}_{\text{Mod}} \tag{16}$$

$$Y_3^h = \beta^h_{g=>2} - \frac{\beta^h_{g=[\text{-}1,0]} + \beta^h_{g=[0,1]}}{2} = \bar{\beta}_{\text{ExtPos}} - \bar{\beta}_{\text{Mod}} \tag{17}$$

$$Y_4^h = \beta^h_{g=<\text{-}2} - \beta^h_{g=>2} = \bar{\beta}_{\text{ExtNeg}} - \bar{\beta}_{\text{ExtPos}} \tag{18}$$

where β^h_g represent the estimated betas from either the daily or cross-sectional regressions discussed in section 3.3. Therefore, the variable Ys are either daily time-series, or cross-sectional series. Moreover, each Y has five versions ($h = 1, \ldots, 5$) corresponding to the horizon at which the hypotheses are evaluated. Equation (15) measures the strength of RH investors reaction to all extreme returns compared to moderate returns, and is directly related to *H1*. Equation (18) subtracts the strength of reaction to large positive returns from the strength of reaction to large negative returns. This is one way to measure asymmetry of reaction and is directly related to *H2*. Equations (16) and (17) evaluate independently the reaction to large negative and large positive returns, respectively, and will be useful for both *H1* and *H2*. For *H3*, which relates to the speed of reaction to extreme returns, we define the following proxies:

$$Y_5 = Y_1^{h=1} - Y_1^{h=5} = \bar{\beta}_{\text{Ext1h}} - \bar{\beta}_{\text{Ext5h}} \tag{19}$$

$$Y_6 = Y_2^{h=1} - Y_2^{h=5} = \bar{\beta}_{\text{ExtNeg1h}} - \bar{\beta}_{\text{ExtNeg5h}} \tag{20}$$

$$Y_7 = Y_3^{h=1} - Y_3^{h=5} = \bar{\beta}_{\text{ExtPos1h}} - \bar{\beta}_{\text{ExtPos5h}} \tag{21}$$

Equation (19) assesses the speed of reaction to all extreme returns, while equations (20) and (21) measure the speed of reaction to extreme negative returns and extreme positive returns, respectively.

Figure 5 displays these proxy variables in the time-series (*i.e.*, based on the betas from daily regressions). All variables exhibit fluctuations over time, confirming that our three hypotheses can be more or less true on different days of the period. For example, with respect to our asymmetric reaction hypothesis (*H2*), we notice in Panel D that our proxy sometimes moves into negative territory (*e.g.*, in early 2020). For these days, the interpretation of *H2* is reversed: the average reaction to extreme positive returns is stronger than the average reaction to extreme negative returns. The higher speed of reaction of RH investors to past extreme negative returns (*H3*, Panel F), while valid for most of the period, can also disappear in certain days (see *e.g.*, early days of July, 2018). It is also interesting to see that, for most proxies, a change seems to occur around the COVID-19 pandemic announcement (March, 2020). Therefore, we will formally test if such an event triggered a structural change in the way RH investors respond to past returns.

Figure 6 displays these proxy variables in the cross-section (*i.e.*, based on the betas from cross-sectional regressions). Results can also vary greatly across companies. For example, while for most companies, as stated in *H1*, RH investors respond to large negative returns by opening new positions within an hour, the left tail of the distribution in Panel B shows that for certain companies they behave in the opposite manner and actually respond by closing their existing positions. Panel D also suggests that the asymmetric reaction observed on aggregate does not manifest ($x \approx 0$) or is reversed ($x < 0$) for a relatively large number of companies.[15]

4.1.2. Time-series factors and firms characteristics

We test for three time-series factors, the level of the VIX, the market excess return (MKT), and the sentiment index from the AAII survey (SENT). The latter is a weekly series, that we tranform into daily frequency by linear interpolation. Panel A of Table 4 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 null (0.03%). The sentiment index has an average of 33% bullish investors over the period, ranging from 20% to 46%. Our firm characteristics variables are defined in Table 3. We choose these characteristics as we assume they might impact RH investors' attention. For example, RH investors might pay more attention to bigger or smaller firms (SIZE). We also attempt to incorporate the investing style (value vs growth) in the analysis by including the Book-to-Market (BM) factor. For the potential impact of firm's financial health, we include the Debt-to-Asset (DA) variable. The other variables are market-related characteristics. Among them, we use trading volume (VOL), liquidity level (BIDASK), firm's systematic risk (BETA) and idiosyncratic risk (IVOL). Finally, we also include the sector of the firm (SEC). Summary statistics for firm characteristics and the number of firms per sector are presented in Panels B and C of Table 4, respectively.[16]

4.2. Time-series regressions

4.2.1. Specifications

Our main specification is as follows:

$$Y_t = \lambda_0 + \lambda_{VIX} VIX + \lambda_{MKT} MKT + \lambda_{SENT} SENT_t + e_t \,, \tag{22}$$

where Y_t is one the proxy variables defined above and VIX, MKT and $SENT$ are the time-series factors. In addition, we develop two specifications to assess the impact of the

[15] To keep results concise, Figures 5 and 6 only show Y_1^h to Y_4^h at the one-hour horizon. For complete results including horizon of two to five hours, refer to Figure XXX in appendix.

[16] Note that the sector repartition is quite uneven (*e.g.*, 429 firms out of 1,426 are from the Health Care sector, and only 9 are from the Real Estate sector. For this reason, the industry-level analysis is not developed in depth.

COVID-19 pandemic announcement on our results. Using March 11, 2020 as the cut-off date (the day the World Health Organization declared the COVID-19 outbreak as a pandemic), we introduce an indicator variable for the period before and after the announcement. We run the dependent variable on this indicator variable in a first specification and test for difference between means. The second specification adds the time-series factor as controls to the first specification.

4.2.2. Results

Table 6 presents the regressions results for dependent variables related to *H1* and *H2*. For the first dependent variable and the first specification $(Y_1, A(1))$, the VIX factor is negative and significant. This indicates that RH investors' reaction to extreme returns is stronger on days of low volatility. When we split the extreme returns into negative and positive ($[Y_2, B(1)]$ and $[Y_3, C(1)]$), the VIX coefficient remains negative and significant, suggesting that this factor has the same impact on both sides of the extreme returns. The MKT factor, however, is not significantly different from zero for the dependent variable representing all extreme returns $(Y_1, A(1))$. Interestingly, when we look at each side of the extreme returns, this factor is actually *positive* (and significant at the 10% level) for large negative returns and *negative* (and significant at the 1% level) for large positive returns. Therefore, it seems that RH investors respond stronger (weaker) to large negative (positive) returns on days of market growth (downturn). The $SENT$ coefficient is strongly positive and significant for the pair $(Y_1, A(1))$. Because it is not significantly different from zero for the pair $(Y_3, C(1))$, this result is mainly driven by the negative side of the extreme returns. It means that, during days of optimism, RH investors tend to respond stronger to large negative returns, but not necessarily to large positive returns. The fourth dependent variable, Y_4, which is a measure of asymmetry, do not seem to depend on market volatility or sentiment. It is however positively correlated with market excess returns. Hence, according to this measure, RH investors seem to respond stronger (weaker) to large negative return compared to large positive return (*i.e.* their reaction is more asymmetrical) on days of market growth (decline). With regards to the two additional specifications assessing the impact of the COVID-19 pandemic, the first three dependent variables have higher means during the pre-COVID period than during the post-COVID period (see pairs $[Y_1, A(2)]$, $[Y_2, B(2)]$, $[Y_3, C(2)]$). The differences between means are significant at the 1%, 4% and 8% level, respectively. This appears consistent with the negative VIX coefficient discussed above: during market turmoils (post-COVID), RH investors react less to extreme returns. When the market is more calm (pre-COVID), RH investors react more to extreme returns. When we add the time-series factors to the COVID dummies (pairs $[Y_1, A(3)]$, $[Y_2, B(3)]$, $[Y_3, C(3)]$), the difference is still significant for all extreme returns and negative extreme returns, but not for positive extreme returns (p-value of 0.379). Interestingly, for Y_1 and Y_2, the means pre-COVID are now *lower* than the means post-COVID. This can be explained by the addition of the VIX factor, whose coefficient is negative and significant, and therefore captures the effect. With regards to our last dependent variable (pairs $[Y_4, D(2)]$ and $[Y_4, D(3)]$), the means pre- and post-COVID are not statistically different.

Table 7 presents the regressions results for the dependent variables related to *H3*. As we found in the previous section that *H3* is only verified for past large negative returns, we emphasize the results related to the corresponding proxy variable, Y_6. The VIX has a negative impact on RH investors' speed of reaction to large negative returns ($[Y_6, F(1)]$). The MKT and $SENT$ factors have a positive impact, but only at 10% level. Therefore, RH investors tend to respond faster to stocks exhibiting past negative returns on days of low volatility, market rise, and investors' optimism. For the specification assessing the effect of COVID ($[Y_6, F(2)]$), the dependent variable means during the pre- and post-COVID periods are not statistically different (p-value of 0.31). When we control for the time-series factors however ($[Y_6, F(3)]$), the pre-COVID mean is lower than its post-COVID counterpart, and the difference becomes significant. It is therefore difficult to determine a structural change in the speed of reaction to large negative returns before and after the COVID-19 pandemic announcement. This last specification also highlights the importance of the VIX and $SENT$ factors which are still able to capture some of the variations in the dependent variable after the inclusion of the COVID dummies.

4.3. Cross-sectional regressions

4.3.1. Specifications

We first examine heterogeneity at the industry-level with the following specification:

$$Y_i = \sum_{s=1}^{11} \lambda_s SEC_i + e_i \, , \tag{23}$$

where Y_i is one the proxy variables defined above, s represents one of the 11 GICS sectors defined in Table 3, and SEC_i is a categorical variable equal to one if the company i belongs to sector s, and zero otherwise.

Then, we run our main analyses which investigate variations at the stock-level:

$$\begin{aligned} Y_i = \lambda_0 &+ \lambda_{SIZE} SIZE_i + \lambda_{BM} BM_i + \lambda_{DA} DA_i + \lambda_{VOL} VOL_i \\ &+ \lambda_{BIDASK} BIDASK_i + \lambda_{BETA} BETA_i + \lambda_{IVOL} IVOL_i + e_i \, , \end{aligned} \tag{24}$$

where the regressors are now the firm characteristics defined in Table 3. To control for industry effect, we also run a version of (24) which includes the SEC categorical variables.

4.3.2. Results

Table 8 presents estimation results from regression (23). RH investors' reaction to past extreme returns (Y_1) seems to vary from one industry to another. The stronger average reaction concerns stocks from the Energy sector (8.38BP), and the lower average reaction concerns stocks from the Utilities sector (3.25BP). When we look at the reaction to large negative and positive returns separately (Y_2 and Y_3), these rankings change. For example, the consumption sectors(Consumption Discretionary, Consumption Staples) are the ones for which RH investors respond the most to past negative returns (12.77BP and 12.19BP, respectively). The asymmetry of reaction to past extreme returns (Y_4) is also different across industries. Firms from the Financials (Health Care) sector are the ones for which RH investors reaction is the most (less) asymmetrical. Dependent variables related to *H3* seem to have less variation among sectors. For example, Y_5 has an average between 0.8BP and 2.8BP for all sectors.

We now turn to our main cross-sectional regressions results. Table 9 shows estimation results for dependent variables related to *H1* and *H2*. For the first dependent variable and the specification without the sector categorical variable ($Y_1, A(1)$), the coefficients associated to DA, VOL, $BIDASK$, $BETA$ and $IVOL$ are positive and statistically different from zero. Therefore, it suggests that stocks with low financial health, high trading volume, low liquidity, and high risk (both systematic and specific) are positively related to the strength of RH investors' reaction to past extreme returns. The factor that matters are not necessarily the same when we separate the reaction to negative and positive returns. Indeed, on the negative side ($Y_2, B(1)$), the $SIZE$ and BM factors become significant (and positive), and the VOL factor is no longer different from zero. Interestingly, on the positive side ($Y_3, C(1)$), the $SIZE$ coefficient is also significant but with a negative sign, the VOL coefficient is different from zero but the BM, DA and $BETA$ estimates no longer are. Putting these results together suggest that firm's size has an opposite effect on the reaction to past negative and positive returns, respectively: RH investors reaction is stronger (weaker) on big firm exhibiting past large negative (positive) returns. On aggregate these two effects cancel out. Besides, the impact of the financial health and systematic risk characteristics to the reaction to all extreme returns is mainly driven by the negative returns side. Similarly, the impact of trading volume to the reaction to all extreme returns is mainly driven by the positive returns side. With regards to the fourth dependent variable measuring asymmetry ($Y_4, D(1)$), our results suggest that all firms' characteristics have a significant impact on its variations in the cross-section. Hence, the asymmetry of reaction is stronger for firms that (i) are bigger, (ii) are value stocks (*i.e.*, higher book-to-market ratio), (iii) have low financial health, (iv) have lower trading volume, (iv) are more liquid, (v) have higher systematic risk and (vi) have higher specific risk. Finally, we note that adding the sectors control variables ($[Y_1, A(2)]$, $[Y_2, B(2)]$, $[Y_3, C(2)]$, $[Y_4, D(2)]$) does not impact the results.

Table 10 presents the regressions results for the dependent variables related to *H3*. As we noticed earlier, *H3* is only verified for extreme negative returns, *i.e.* RH investors reaction to extreme negative return is stronger during the first hours. Therefore, we take a particular interest in the dependent variable Y_6. The speed of reaction to stocks exhibiting large negative returns ($[Y_6, F(1)$) is positively impacted by the $SIZE$, DA, $BETA$, and $IVOL$ factors. Therefore, RH investors tend to respond faster to firms exhibiting large past negative returns that (i) are bigger in size, (ii) have low financial health, (iii) have higher systematic risk, and (iv) have higher specific risk. These results hold when firms' sectors variables are included ([$Y_6, F(2)$).

4.4. Results Summary

4.4.1. Time-Series

H1 RH investors' reaction to stocks with past extreme returns is stronger (weaker) on days of high (low) market volatility and investors' optimism (pessimism). It is not significantly related to market excess returns.

H2 RH investors' asymmetry of reaction to extreme returns is stronger (weaker) on days when market is up (down). Furthermore, when we separate extreme returns into negative and positive returns, we see that this exacerbated asymmetry comes from both sides of the extreme returns: on days when the market is rising, RH investors tend to react stronger to stocks with large negative returns and weaker to stocks with large positive returns. This asymmetry is not significantly related to market volatility or investors sentiment.

H3 RH investors' speed of reaction to extreme negative returns is higher (lower) on days of low (high) market volatility, on days of market rise (decline), and on days of investors' optimism (pessimism).

4.4.2. Cross-Section

H1 RH investors' reaction to stocks with past extreme returns is stronger (weaker) for firms with bad (good) financial health, high (low) trading volume, low (high) liquidity, high (low) systematic and high (low) specific risk;

H2 RH investors' asymmetry of reaction to extreme returns is stronger (weaker) for larger (smaller) firms, value (growth) stocks, firms with bad (good) financial health, firms with low (high) trading volume, firms with high (low) liquidity, firms with high (low) systematic risk and firms with high (low) specific risk. In addition, firm's size is positively related to the RH investors' response to extreme negative returns, but negatively related to the RH investors' response to extreme positive returns.

H3 RH investors' speed of reaction to extreme negative returns is higher (lower) for larger (smaller) firms, firms with bad (good) financial health, firms with high (low) systematic risk and firms with high (low) specific risk.

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.*

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.

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.

Grinblatt, M., Keloharju, M., Linnainmaa, J., 2011. IQ and stock market participation. Journal of Finance 66, 2121–2164. doi:10.1111/j.1540-6261.2011.01701.x.

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.

Shive, S., 2010. An epidemic model of investor behavior. Journal of Financial and Quantitative Analysis 45, 169–198. doi:10.1017/S0022109009990470.

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.

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 (main variables)
Panels A and B present summary statistics across stock-hour observations for our Robinhood investors' trading behavior ($\Delta RH_{i,t,h}$) and adjusted returns proxies ($\tilde{r}_{i,t,h}$), respectively. The number of observations is expressed in million. Statistics from panel A are expressed in basis point, except for N, T and #, which represent the number of observations, days, and companies, respectively.

Panel A: Robinhood users change										
	Av	Std	Min	25th	50th	75th	Max	N	T	#
Intraday	4.41	122.71	-6937.25	-2.83	0.00	2.14	59,602.68	6,793,214	527	2,853
Overnight	4.09	59.88	-15126.30	-3.44	0.00	5.37	15,040.79	1,238,281	526	2,853
All	4.36	115.28	-15126.30	-3.05	0.00	3.31	59,602.68	8,031,495	527	2,853
Panel B: Adjusted returns										
	Av	Std	Min	25th	50th	75th	Max	N	T	#
Intraday	-0.01	1.00	-30.33	-0.42	0.00	0.41	44.42	6,793,214	527	2,853
Overnight	0.01	1.00	-20.80	-0.34	0.01	0.37	20.61	1,238,281	526	2,853
All	-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 β_g estimates from regression in equation (12) for horizon of one to five hours. All five regressions are based on the full panel and are estimated by standard OLS. Estimates are expressed in basis point. 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	h=1	h=2	h=3	h=4	h=5
<-2	7.95***	6.07***	4.98***	4.2***	3.57***
	(94.5)	(73.63)	(61.89)	(53.67)	(46.06)
[-2,-1]	2.54***	2.1***	1.62***	1.29***	0.99***
	(66.2)	(55.91)	(43.77)	(35.29)	(27.02)
[-1,0]	0.31***	0.28***	0.21***	0.17***	0.15***
	(19.54)	(17.48)	(12.9)	(10.22)	(8.82)
[0,1]	-0.2***	-0.16***	-0.05***	0.02	0.08***
	(-13.2)	(-10.15)	(-2.91)	(1.16)	(4.79)
[1,2]	0.03	0.28***	0.47***	0.62***	0.68***
	(0.66)	(7.32)	(12.34)	(16.48)	(17.94)
>2	2.71***	2.61***	2.68***	2.7***	2.78***
	(30.4)	(29.92)	(31.74)	(33.22)	(34.41)
R-squared	0.014	0.015	0.014	0.014	0.014
Nobs	8017230	8017230	8017230	8017230	8017230

Table 3: 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 4: Summary Statistics (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 3.

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	1426
BM	0.44	1.08	-19.74	0.13	0.30	0.63	11.81	1426
DA	0.32	0.30	0.00	0.10	0.28	0.44	5.40	1426
VOL	13.65	1.31	9.59	12.68	13.66	14.51	18.11	1426
BIDASK(%)	0.32	0.50	0.01	0.04	0.10	0.34	3.98	1426
BETA	1.19	0.44	-0.73	0.92	1.18	1.45	3.12	1426
IVOL(%)	3.98	2.63	0.81	2.09	3.34	5.05	27.84	1426

Table 5: Sectors' distribution

	GICS sector	N
1	Energy	96
2	Materials	53
3	Industry	134
4	Consumer discretionary	221
5	Consumer staple	62
6	Health	429
7	Financials	84
8	Information Technology	233
9	Telecommunications	74
10	Utilities	31
11	Real Estate	9

Table 6: Time-series Regression results, *H1* and *H2*

Each panel corresponds to the proxy variable defined in equations (15) to (18), respectively. *Pr(>F)* corresponds to the p-value of the F-test evaluating the equality of COV(0) and COV(1). The cutoff date defining the period before and after COVID is March 11, 2020. 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. Returns ($Y_1^{h=1}$)			Ext. Neg. Returns ($Y_2^{h=1}$)			Ext. Pos. Returns ($Y_3^{h=1}$)			Asymmetry ($Y_4^{h=1}$)		
	A(1)	A(2)	A(3)	B(1)	B(2)	B(3)	C(1)	C(2)	C(3)	D(1)	D(2)	D(3)
Intercept	4.33***			4.64***			4.02*			0.62		
	(3.18)			(3.26)			(1.71)			(0.23)		
COV(0)		6.92***	4.21***		9.42***	4.47***		4.42***	3.95*		5***	0.53
		(17.24)	(2.95)		(20.05)	(2.73)		(7.45)	(1.7)		(7.07)	(0.19)
COV(1)		5.24***	6***		8.03***	6.98***		2.44***	5.02*		5.59***	1.96
		(14.03)	(3.65)		(16.42)	(3.72)		(2.61)	(1.79)		(4.32)	(0.57)
VIX	-0.09***		-0.14***	-0.08***		-0.14***	-0.1***		-0.13***	0.02		-0.02
	(-6.22)		(-3.97)	(-3.74)		(-3.94)	(-4.31)		(-2.89)	(0.59)		(-0.35)
MKT	-0.04		-0.12**	0.31*		0.2	-0.4***		-0.45**	0.71**		0.64**
	(-0.72)		(-2.01)	(1.71)		(1.25)	(-2.63)		(-2.4)	(2.27)		(2)
SENT	12.44***		14.53***	18.56***		21.49***	6.33		7.58	12.23		13.91
	(2.93)		(3.47)	(4.04)		(4.41)	(0.8)		(0.98)	(1.25)		(1.41)
Pr(>F)		0.002	0.016		0.037	0.015		0.076	0.379		0.692	0.4
Adj.R^2	0.175	0.816	0.846	0.133	0.809	0.836	0.078	0.451	0.481	0.034	0.388	0.41
Nobs	526	526	526	526	526	526	526	526	526	526	526	526

Table 7: Time-series Regression results, *H3*

Each panel corresponds to the proxy variable defined in equations (19) to (21), respectively. *Pr(>F)* corresponds to the p-value of the F-test evaluating the equality of COV(0) and COV(1). The cutoff date defining the period before and after COVID is March 11, 2020. 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Speed Ext. Ret. (Y_5)			Speed Ext. Neg. Ret. (Y_6)			Speed Ext. Pos. Ret. (Y_7)		
	E(1)	E(2)	E(3)	F(1)	F(2)	F(3)	G(1)	G(2)	G(3)
Intercept	4.91***			5.7***			4.13***		
	(3.92)			(3.65)			(2.59)		
COV(0)		5.14***	4.82***		7.07***	5.5***		3.22***	4.15***
		(17.03)	(3.87)		(18.47)	(3.54)		(8.34)	(2.6)
COV(1)		3.53***	6.13***		6.46***	8.4***		0.61	3.85**
		(6.73)	(4.86)		(12.61)	(5.41)		(0.73)	(2.08)
VIX	-0.1***		-0.13***	-0.08***		-0.15***	-0.11***		-0.11***
	(-8.92)		(-5.7)	(-5.12)		(-5.91)	(-5.75)		(-3.77)
MKT	-0.02		-0.08	0.34*		0.21	-0.37**		-0.36**
	(-0.29)		(-1.24)	(1.89)		(1.31)	(-2.35)		(-2.1)
SENT	5.61		7.12*	8.49*		11.87**	2.72		2.38
	(1.5)		(1.92)	(1.75)		(2.5)	(0.53)		(0.46)
Pr(>F)		0.006	0.01		0.31	0		0.005	0.732
Adj.R^2	0.12	0.692	0.722	0.071	0.696	0.725	0.078	0.291	0.312
Nobs	526	526	526	526	526	526	526	526	526

Table 8: Cross-section Regressions results, industry level

This table shows the λ_s estimates from regression ((23)), estimated by standard OLS. Estimates are expressed in basis point. The associated t-statistics are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	H1 and H2				*H3*		
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
EN	8.38***	11.75***	5.02***	6.73***	2.71***	6.13***	-0.71
	(18.7)	(19.7)	(7.26)	(7.24)	(7.5)	(12.89)	(-1.22)
MAT	3.69***	7.52***	-0.13	7.65***	1.17**	4.41***	-2.08***
	(6.12)	(9.36)	(-0.14)	(6.11)	(2.4)	(6.89)	(-2.66)
IND	5.87***	9.97***	1.77***	8.2***	1.96***	5.1***	-1.18**
	(15.47)	(19.75)	(3.03)	(10.42)	(6.39)	(12.66)	(-2.41)
CDISC	7.52***	12.77***	2.28***	10.49***	2.44***	6.64***	-1.75***
	(25.46)	(32.48)	(4.99)	(17.12)	(10.24)	(21.17)	(-4.58)
CSTA	5.35***	9.05***	1.65*	7.4***	2.1***	5.37***	-1.17
	(9.59)	(12.19)	(1.92)	(6.4)	(4.66)	(9.08)	(-1.62)
HLTH	6.48***	7.29***	5.67***	1.62***	3.77***	4.89***	2.65***
	(30.56)	(25.83)	(17.34)	(3.67)	(22.02)	(21.74)	(9.63)
FIN	5.56***	11.09***	0.02	11.07***	1.62***	5.93***	-2.69***
	(11.6)	(17.39)	(0.03)	(11.14)	(4.19)	(11.66)	(-4.33)
IT	4.51***	7.95***	1.08**	6.87***	2.36***	5.24***	-0.51
	(15.68)	(20.75)	(2.43)	(11.51)	(10.17)	(17.15)	(-1.37)
TEL	4.94***	8.48***	1.4*	7.07***	1.98***	5.08***	-1.13*
	(9.67)	(12.48)	(1.78)	(6.68)	(4.79)	(9.37)	(-1.71)
UT	3.25***	7.47***	-0.97	8.44***	0.81	4.26***	-2.64***
	(4.12)	(7.12)	(-0.8)	(5.16)	(1.27)	(5.08)	(-2.59)
RE	7.44***	11.93***	2.94	9***	2.8**	5.91***	-0.3
	(5.08)	(6.13)	(1.3)	(2.96)	(2.37)	(3.8)	(-0.16)
Adj.R^2	0.663	0.722	0.215	0.379	0.375	0.574	0.095
Nobs	1426	1426	1426	1426	1426	1426	1426

Table 9: Cross-section Regression results, *H1* and *H2*

Each panel corresponds to the proxy variable defined in equations (15) to (18), respectively. *INDCTRL* specifies whether the industry dummy variables are included in the regression. All specifications are estimated by standard OLS. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

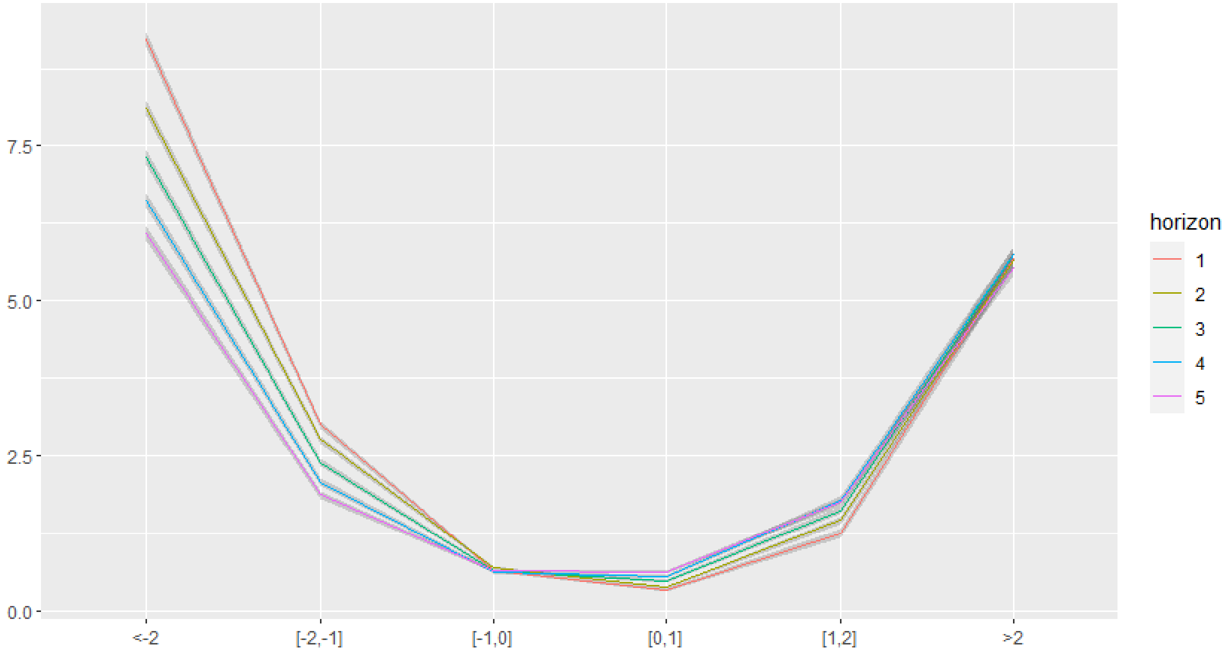
	Ext. Returns ($Y_1^{h=1}$)		Ext. Neg. Returns ($Y_2^{h=1}$)		Ext. Pos. Returns ($Y_3^{h=1}$)		Asymmetry ($Y_4^{h=1}$)	
	A(1)	A(2)	B(1)	B(2)	C(1)	C(2)	D(1)	D(2)
Intercept	-11.69***	-10.23***	-30.49***	-28.92***	7.1***	8.47***	-37.59***	-37.39***
	(-6.26)	(-5.23)	(-11.89)	(-10.89)	(2.77)	(3.11)	(-10.69)	(-10.11)
SIZE	0.04	0.05	1.29***	1.33***	-1.21***	-1.23***	2.5***	2.57***
	(0.46)	(0.55)	(10.05)	(10.63)	(-9.37)	(-9.56)	(14.16)	(14.68)
BM	0.1	0.01	0.33**	0.2	-0.13	-0.17	0.47**	0.38*
	(0.96)	(0.13)	(2.36)	(1.43)	(-0.96)	(-1.2)	(2.42)	(1.9)
DA	1.17***	0.67*	2.65***	1.6***	-0.31	-0.26	2.96***	1.86***
	(3.22)	(1.83)	(5.31)	(3.21)	(-0.62)	(-0.5)	(4.32)	(2.67)
VOL	0.81***	0.81***	0.24	0.18	1.39***	1.45***	-1.14***	-1.27***
	(6.6)	(6.58)	(1.44)	(1.05)	(8.16)	(8.41)	(-4.91)	(-5.45)
BIDASK(%)	2.9***	2.94***	1.5***	1.59***	4.31***	4.29***	-2.81***	-2.7***
	(8.84)	(9.08)	(3.32)	(3.62)	(9.54)	(9.5)	(-4.53)	(-4.4)
BETA	1.82***	1.39***	4.06***	3.39***	-0.43	-0.62	4.49***	4.02***
	(7.06)	(5.04)	(11.49)	(9.08)	(-1.22)	(-1.62)	(9.26)	(7.71)
IVOL(%)	0.55***	0.55***	0.67***	0.75***	0.43***	0.35***	0.23**	0.4***
	(9.07)	(8.84)	(8.01)	(8.89)	(5.2)	(4.03)	(2.04)	(3.41)
INDCTRL	NO	YES	NO	YES	NO	YES	NO	YES
Adj.R^2	0.247	0.278	0.227	0.275	0.412	0.418	0.407	0.426
Nobs	1426	1426	1426	1426	1426	1426	1426	1426

Table 10: Cross-section Regression results, *H3*

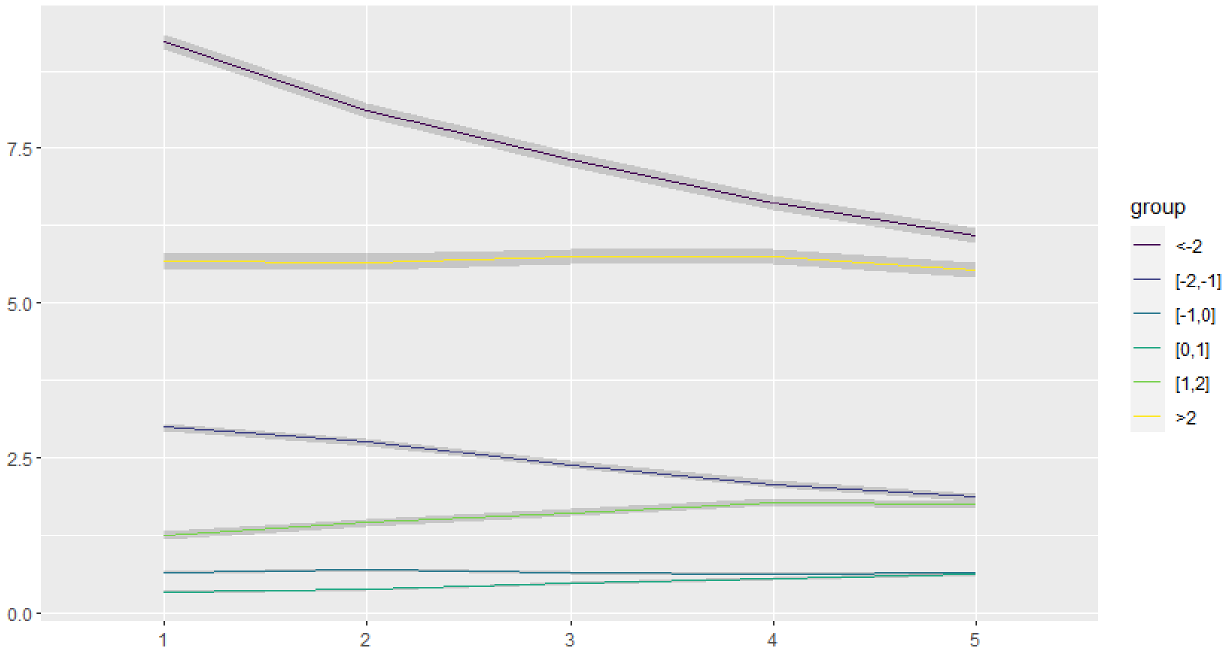
Each panel corresponds to the proxy variable defined in equations (19) to (21), respectively. *IND-CTRL* specifies whether the industry dummy variables are included in the regression. All specifications are estimated by standard OLS. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Speed Ext. Ret. (Y_5)		Speed Ext. Neg. Ret. (Y_6)		Speed Ext. Pos. Ret. (Y_7)	
	E(1)	E(2)	F(1)	F(2)	G(1)	G(2)
Intercept	2.72*	2.31	-7.51***	-7.13***	12.95***	11.75***
	(1.78)	(1.41)	(-3.46)	(-3.09)	(5.56)	(4.75)
SIZE	-0.26***	-0.27***	0.41***	0.43***	-0.93***	-0.96***
	(-3.38)	(-3.46)	(3.77)	(3.93)	(-7.94)	(-8.23)
BM	-0.12	-0.08	0.1	0.09	-0.34***	-0.25*
	(-1.4)	(-0.92)	(0.87)	(0.69)	(-2.65)	(-1.86)
DA	0.07	0.19	1.26***	0.99**	-1.13**	-0.61
	(0.22)	(0.62)	(2.99)	(2.27)	(-2.5)	(-1.3)
VOL	0.31***	0.35***	0.06	0.02	0.56***	0.68***
	(3.06)	(3.38)	(0.4)	(0.15)	(3.65)	(4.32)
BIDASK(%)	1.15***	1.18***	0.47	0.51	1.83***	1.85***
	(4.3)	(4.37)	(1.24)	(1.34)	(4.48)	(4.51)
BETA	-0.43**	-0.49**	1.21***	0.98***	-2.07***	-1.96***
	(-2.06)	(-2.13)	(4.03)	(3.02)	(-6.46)	(-5.63)
IVOL(%)	0.34***	0.31***	0.35***	0.39***	0.33***	0.24***
	(6.82)	(6.04)	(4.94)	(5.27)	(4.34)	(3.05)
INDCTRL	NO	YES	NO	YES	NO	YES
Adj.R^2	0.205	0.207	0.042	0.048	0.319	0.329
Nobs	1426	1426	1426	1426	1426	1426

Figure 1: Average RH investors reaction to stock returns at different horizon
This figure plots the average RH users change (y axis) at the x hours horizon, conditional on the previous adjusted returns being in a certain group level g, with 95% confidence bands. Panel A displays the results using return groups levels (g) for the x-axis. Panel B displays the results using horizon (x) for the x-axis. The RH users change variable is winsorized at the 95% and 5% and is expressed in basis point.



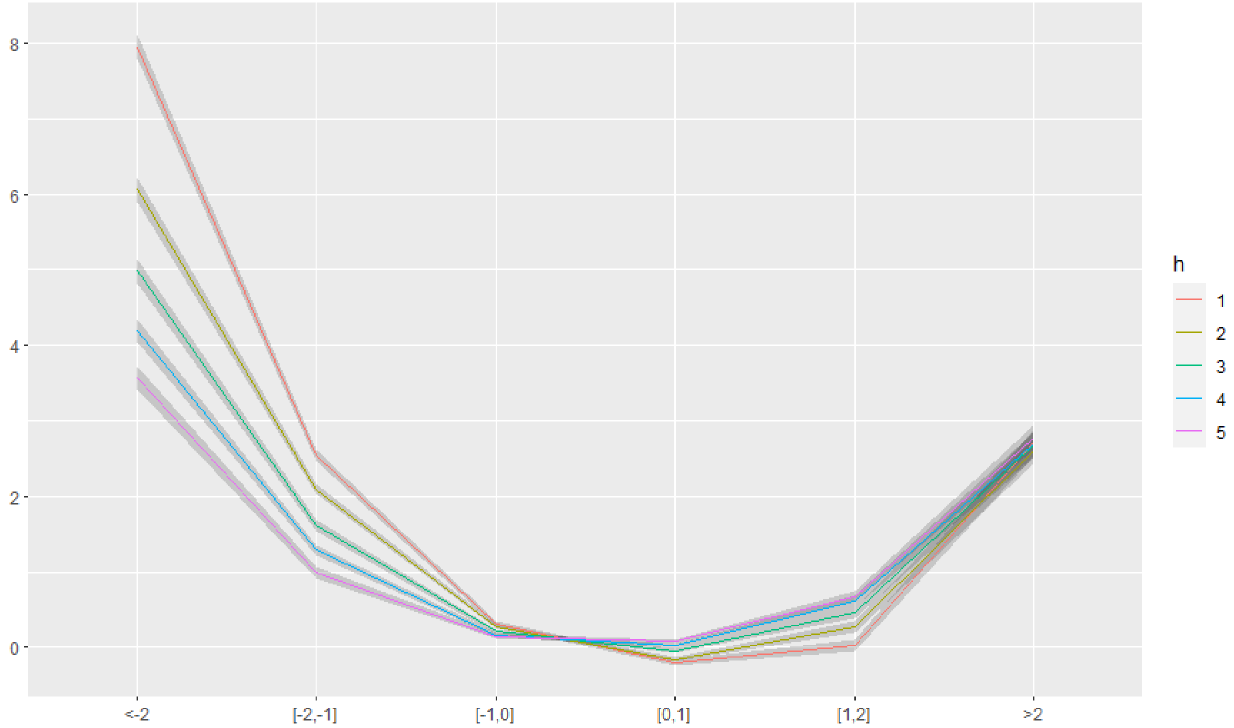
(a) Panel A



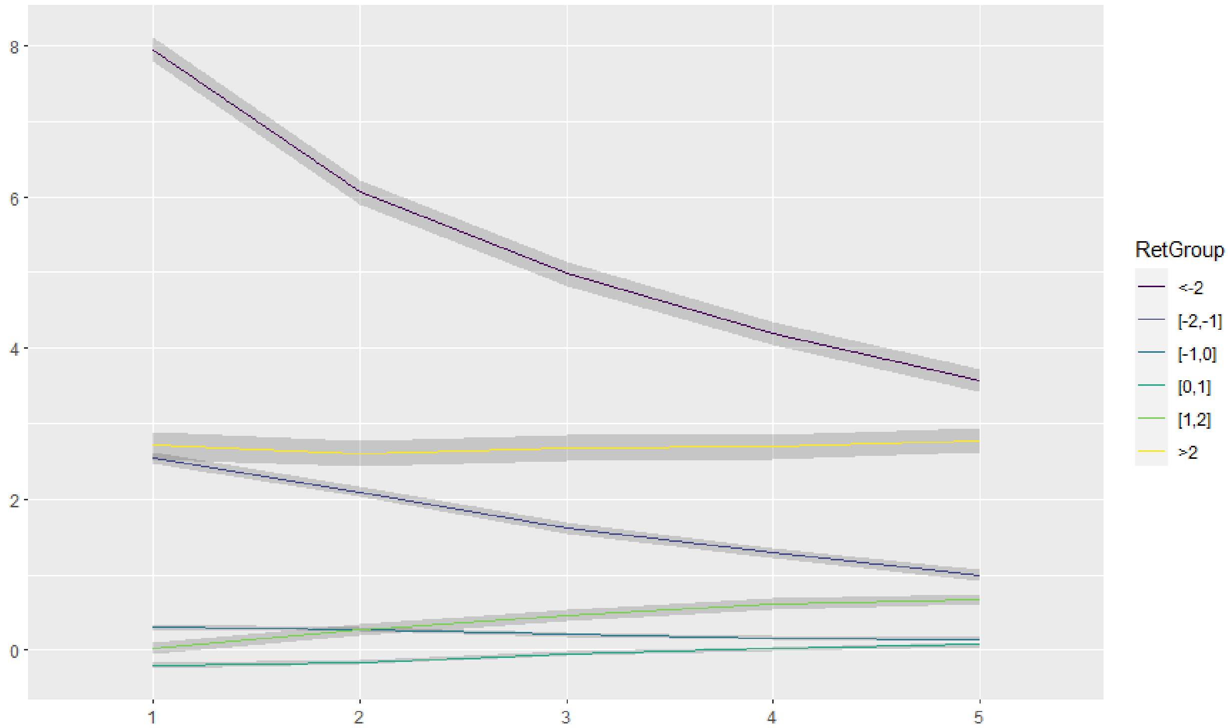
(b) Panel B

Figure 2: Estimates from regressions

This figure plots the β_g coefficients from the regression in equation (12) with the horizon (x) varying from one to five hours, along with 95% confidence bands. Panel A displays the results using return groups levels (g) for the x-axis. Panel B displays the results using horizon (x) for the x-axis. The RH users change variable is winsorized at the 95% and 5% and is expressed in basis point.



(a) Panel A



(b) Panel B

Figure 3: Daily Regressions: sensitivity to extreme previous returns
This figure plots the β_g corresponding to group <-2 and >2 estimated in daily regressions, each with an horizon from one to five hours. The series are expressed in basis point.

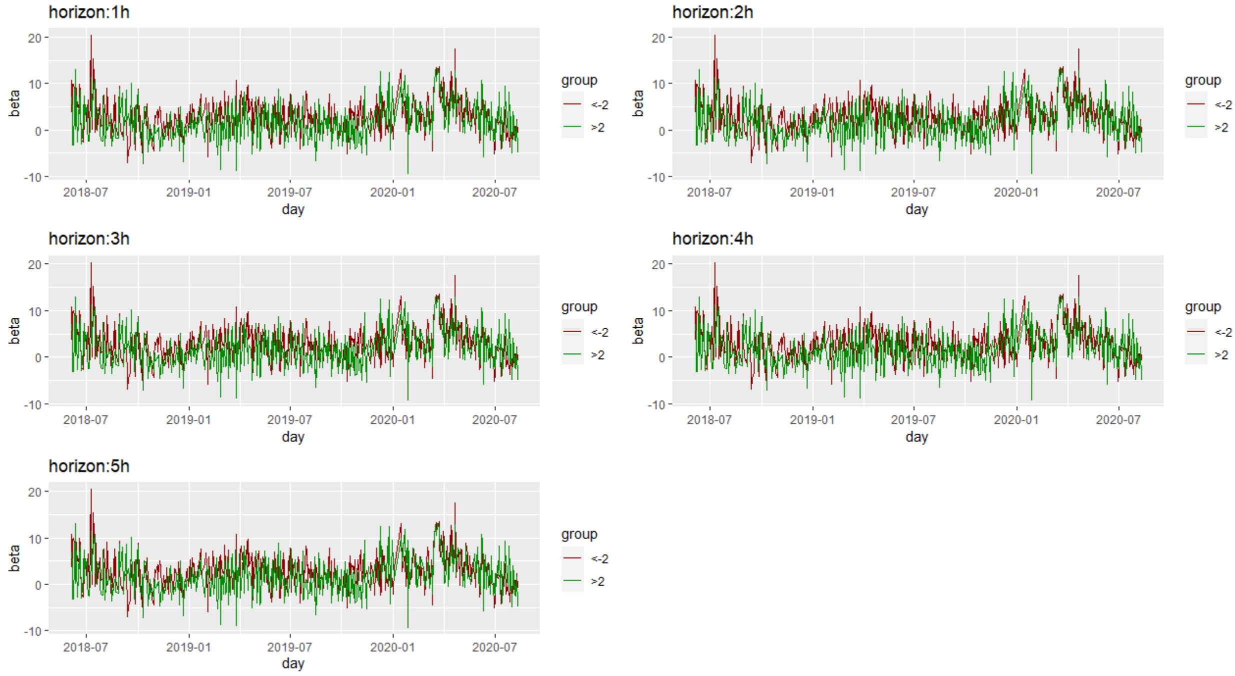


Figure 4: Cross-section Regressions (stock level)
This figure plots the β_g estimated in regressions for each company, with an horizon from one to five hours. The betas are expressed in basis point.

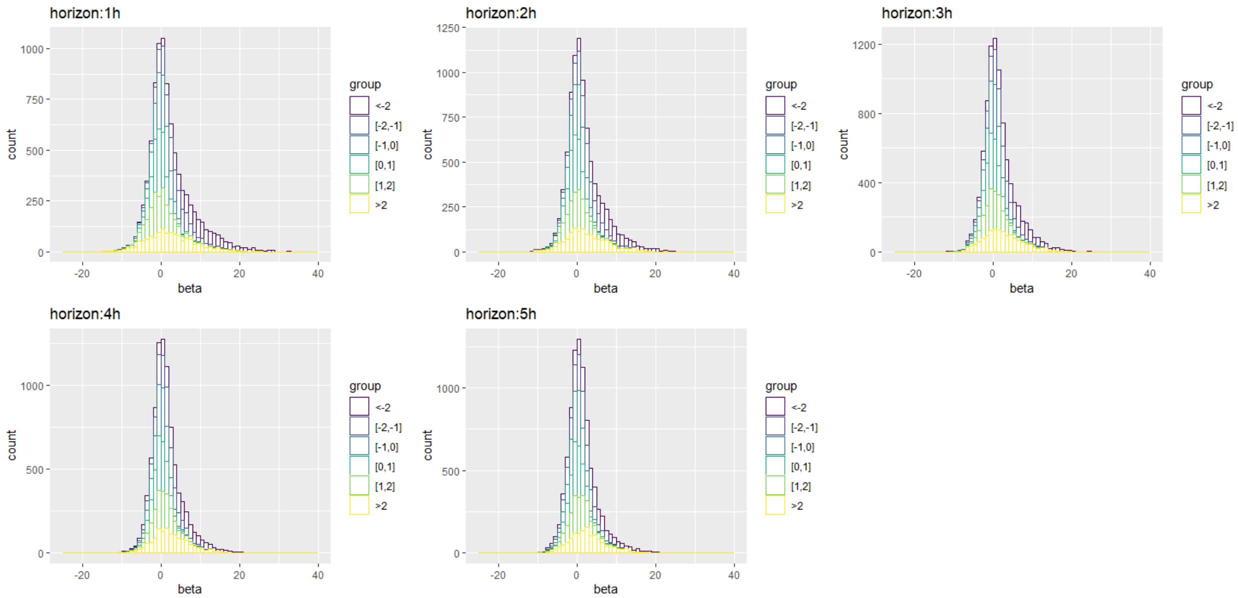
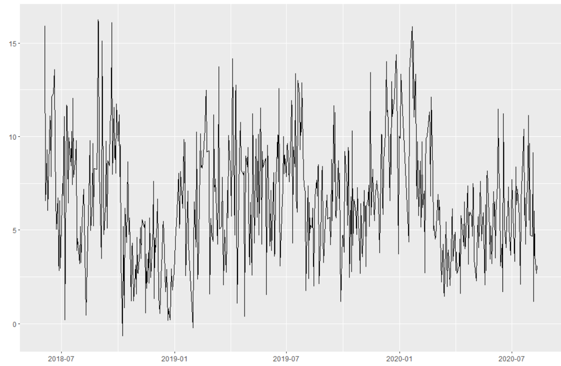
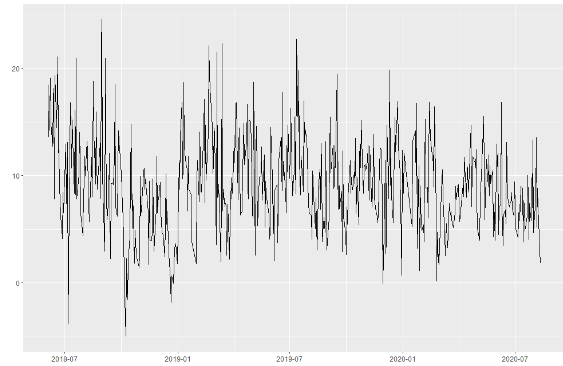


Figure 5: Dependent Variables (time series)

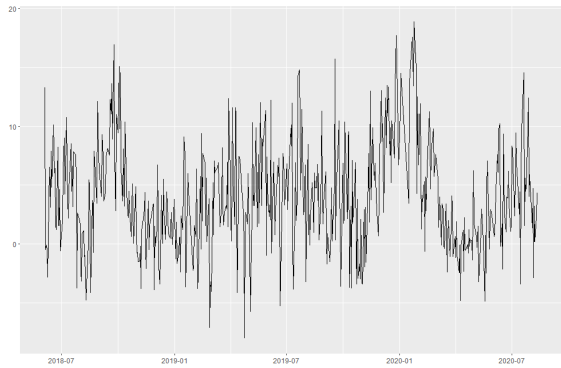
This figure presents the proxy variables constructed with the daily estimated β_g based on regression (12). Panels A to D correspond to the variable defined in equation (15) to (18), respectively, evaluated at the one-hour horizon ($h = 1$). Panels E to G correspond to the variable defined in equation (19) to (21), respectively.



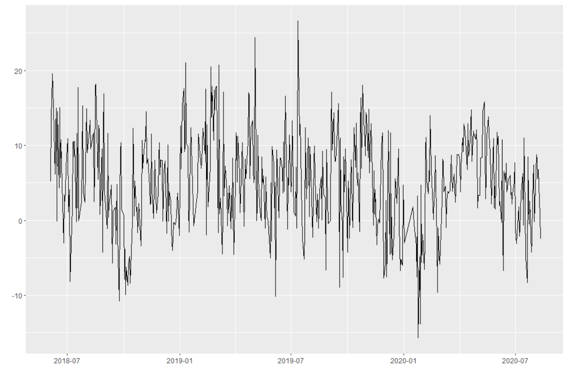
(a) Reaction to ext. returns ($Y_1^{h=1}$)



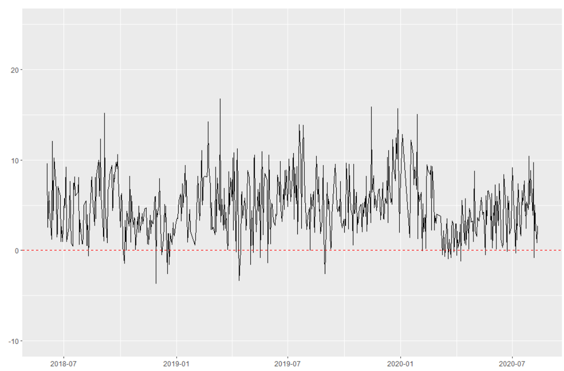
(b) Reaction to ext. negative returns ($Y_2^{h=1}$)



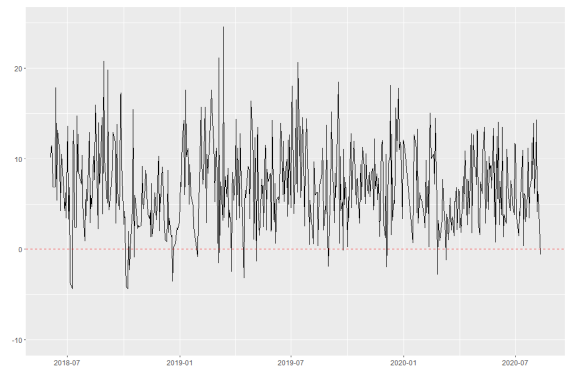
(c) Reaction to ext. positive returns ($Y_3^{h=1}$)



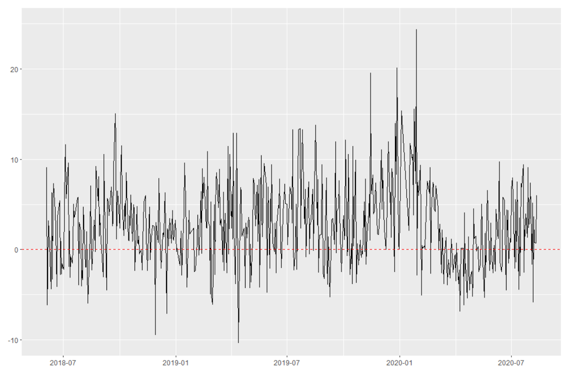
(d) Asymmetry of reaction to ext. returns ($Y_4^{h=1}$)



(e) Speed of reaction to ext. returns (Y_5)

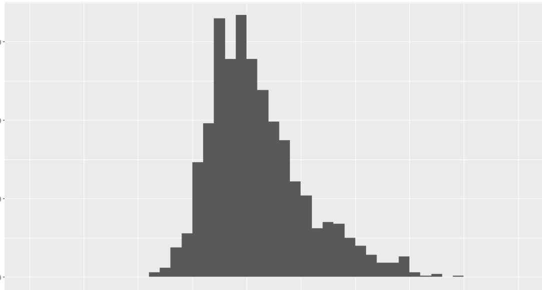


(f) Speed of reaction to ext. negative returns (Y_6)

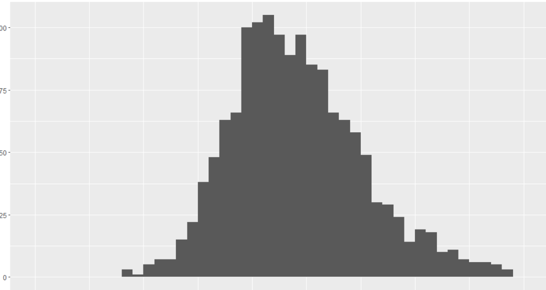


(g) Speed of reaction to ext. positive returns (Y_7)

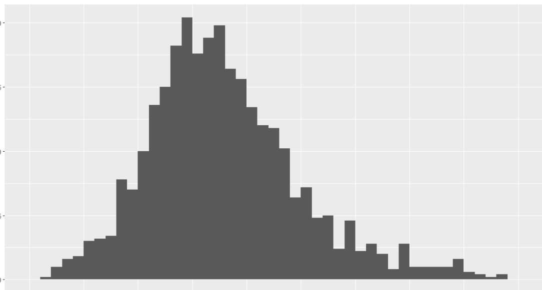
Figure 6: Dependent Variables (cross section)
This figure present the proxy variables based on estimated β_g from the cross sectionnal regressions based on (12). Panels A to D correspond to the variable defined in equation (15) to (18), respectively, evaluated at $h = 1$. Panels E to G correspond to the variable defined in equation (19) to (21), respectively.



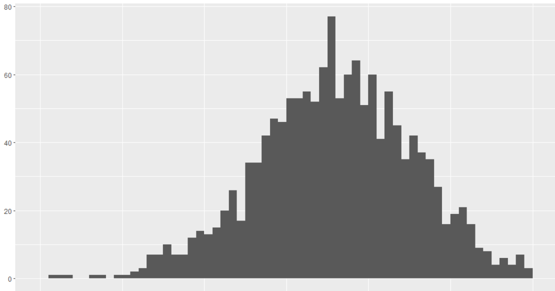
(a) Reaction to ext. returns ($Y_1^{h=1}$)



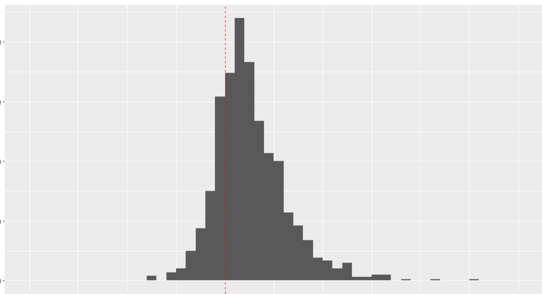
(b) Reaction to ext. negative returns ($Y_2^{h=1}$)



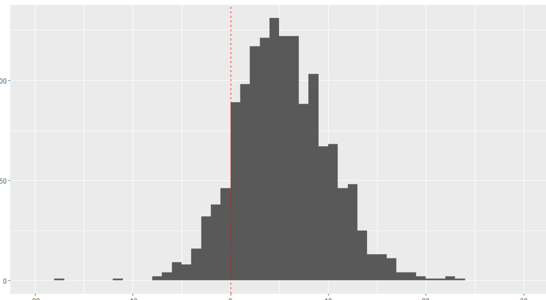
(c) Reaction to ext. positive returns ($Y_3^{h=1}$)



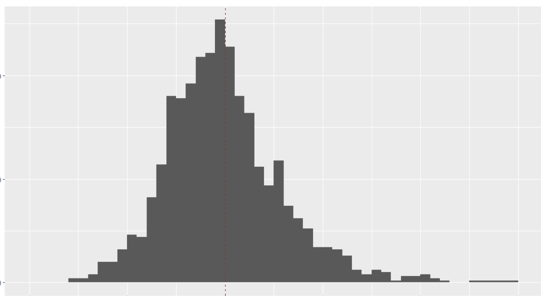
(d) Asymmetry of reaction to ext. returns ($Y_4^{h=1}$)



(e) Speed of reaction to ext. returns (Y_5)



(f) Speed of reaction to ext. negative returns (Y_6)



(g) Speed of reaction to ext. positive returns (Y_7)