
**IMPACT OF INTERNET INFORMATION ON RETAIL INVESTORS’
INVESTMENT DECISION: EVIDENCE FROM ROBINHOOD
INVESTORS AND REDDIT DISCUSSIONS**

Jai Khanna, mentor: Ammon lam

Step by Step School , Noida

DOI: 10.46609/IJSSER.2020.v05i11.026 URL: <https://doi.org/10.46609/IJSSER.2020.v05i11.026>

ABSTRACT

This paper studies how information on the internet affects the stock picking decision of retail investors. We use stock discussion on subReddit r/wallstreetbet to proxy for information that retail investors receive from the internet and look at how stock held by Robinhood investors correlates with the amount of discussion around the stock in Reddit. Our result shows that stocks with a high number of discussions on Reddit are excessively held by Robinhood investors, even after controlling for market capitalization of each stock. Our result sheds light on the process through which retail investors decide which stock to include in their portfolio, and reveals another dimension of irrationality in retail investor’s behavior.

Keywords: Investor, Reddit, stock, Capital, Retail.

INTRODUCTION

A retail investor is an investor who buys and sells stocks, mutual funds, or exchange traded funds. Retail investors buy securities for their own personal accounts and mostly trade in significantly smaller quantities than other investors. Retail investors now account for roughly 20% of stock-market activity on average and nearly one-quarter of trades on peak days.(businessinsider.in) Individually, the investment amount of each household is small, but it becomes an important fraction of the stock market when aggregated.

Studying the behaviour of retail investors is important because their investment mistakes directly affect the welfare of individual investors. By carefully studying the behaviour of retail investors, the government can impose suitable policies and regulations to protect retail investors and potentially direct them to healthy and beneficial investments.

To better our understanding of retail investor trading behavior, we look at how information on the internet affects the stock picking decision of retail investors. In order to do so, we look at a popular retail investors trading platform: Robinhood. Robinhood is a free-trading software that allows investors to trade stocks, futures, exchange-traded funds, and cryptocurrencies without any fee. Robinhood was founded in April 2013 and did extremely well in its first few years. This was because it was one of the only brokers offering free trade. But as more online brokers have started removing trading commissions and fees, Robinhood is becoming less popular among customers.(businessinsider.in) That said, it is still very popular and currently one of the only few brokers that gives investors the opportunity to trade cryptocurrency. Both online and mobile trading are offered through Robinhood. Robinhood is best suited for beginners who want to trade in small quantities, like fractional shares and cryptocurrencies, and require nothing besides seeing what others are investing in terms of their study.

Trading costs for Robinhood are very low and crypto-currency trades can be placed in small quantities, making it an ideal platform. It is sold to customers in the form of an app or website which are fairly easy to use and understand. Robinhood is efficient at getting cash into the market. After closing their positions, all customers have immediate access to deposits to funds. Given these benefits, Robinhood has many fallbacks. There is no commission charged by Robinhood for trades, but the spread we saw for our cryptocurrency transactions was considerably wider than those we saw on other platforms. Though prices update on the Robinhood app and the website, they lag other real-time data providers by several seconds. New investors who are dedicated to improving their trading skills may outgrow the resources provided by Robinhood, especially options traders.

On top of studying retail investor behavior through Robinhood we also found a method to measure the information that retail investors are getting from the internet. For that we have scrapped and analyzed the discussion on Reddit. Started in 2006, Reddit is an American social news aggregation, web content rating, and discussion website. Registered members submit content to the site such as links, text posts, and images, which are then voted up or down by other members. Reddit has over 48 million active users currently. (statistica.com) There are a few subReddits where retail investors gather to discuss the stock market, including r/wallstreetbet, r/stockmarket, and r/stocks. r/wallstreetbet is the biggest subReddit among them, and with the most amount of active discussion. Therefore in this paper we are going to use the discussions in r/wallstreetbet to proxy for retail investor's attention on specific stocks. We reduce the discussion into a frequency of each stock ticker (e.g. APPL, TSLA) being mentioned in r/wallstreetbet.

This ticker discussion frequency is a measure of how often the stock is discussed in Reddit.

(capitalize Reddit)

This paper will attempt to link Reddit discussions and retail investor stock purchase behavior. We will also establish that retail investors do respond to uninformative content on the internet which is an irrational behavior. Our paper will make the reader understand the channel in which investors make investment decisions. This will in turn help make and implement better policies that will be directed towards providing retail investors with better information therefore mitigating the cost from bad investment decisions.

LITERATURE REVIEW

There are several papers available that discuss the information investors use in investment decisions. Riffin and Ahmad (2012) discuss the level of risk investors are willing to absorb and how it will affect their investment strategy: aggressive, moderate or conservative. Furthermore, Divanlioglu and Bagci (2018) give a comprehensive understanding of the factors that influence investor behaviour. The paper suggests that investors make investment decisions rationally by intermittent change from the past and that investment decisions can be taken under the influence of some psychological factors. The paper shows how human behaviors are effective in the functioning of investment decisions and that investment decisions can be made through non-rational behaviors.

In addition to this, there are papers done on how news diffuses into stock prices. Peress(2014) explains how the media is increasingly recognized as key players in financial markets. In this paper the causal impact of media on trading and price formation is investigated by examining national newspaper strikes in several countries. The findings demonstrate that the media contribute to the efficiency of the stock market by improving the dissemination of information among investors and its incorporation into stock prices.

Recent evidence suggests a positive correlation between an increase in internet forum discussions and stock market activity. Using pure message counts, Wysocki (1998) states that variance in daily message posting volume is related to news announcements and earnings. Using computer algorithms, Lavrenko et al. (2000) classify news reports that impact markets, and then exchange this information successfully.

The bullishness of messages is examined by Antweiler and Frank (2004), who find that although web talk does not predict market movements, it is predictive of volatility.

The above three studies reveal that information on the internet affects market outcome such as market movements or volatility. However, none of them looked at actual buying behavior of

investors as a result of information from the internet. Therefore, this motivated us to contribute to this literature by documenting how retail investors use information from the internet as they decide which stocks to buy.

DATA

Robinhood data: The total stock holdings by robinhood users data is obtained by Robintrack. Robintrack keeps track of how many Robinhood users hold a particular stock over time. While this does not equal the actual amount of stock being held in total by robinhood users, it is a good proxy of how many shares is being held by robinhood users if we impose an assumption that each investor that shows up in the data holds one share of the stock.

Reddit (r/wallstreetbet) data: The discussion on r/wallstreetbet is analysed using a python script. We record the frequency of each stock ticker (e.g. APPL, TSLA) and mention it in the comments section of r/wallstreetbet. This tells us how frequently a stock is being discussed in Robinhood. Among the thousands of stocks being traded on Robinhood, half of them were not ever discussed on r/wallstreetbet. For this study we focus our attention on stocks that is being mentioned at least once on r/wallstreetbet

MAIN RESULT

Figure 1a and 1b: Histogram of discussion on Reddit and histogram of stock holdings on Robinhood

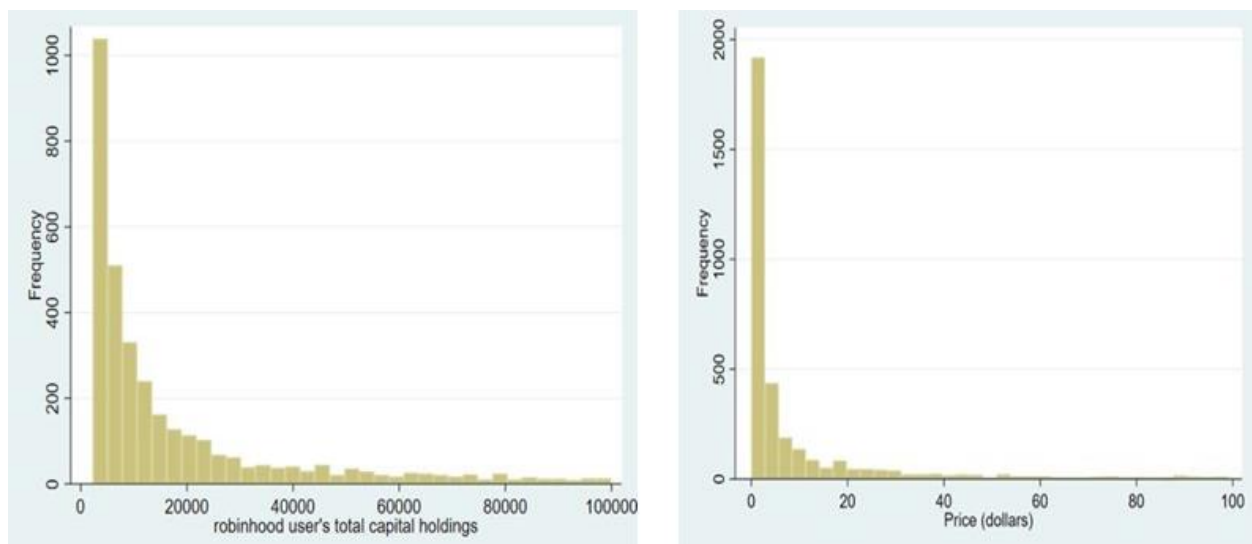


Figure 1a is a histogram showing the frequency of Robinhood user's total capital holdings. The x axis of the graph shows a range of values for robinhood user's total capital holdings and the y axis shows the frequency.

We see a decline in frequency the total robinhood holdings of a stock increases. There is a large amount of stock which robinhood users do not have any holdings.

Figure 1b is a histogram showing the frequency of stocks prices. The x axis of the graph shows the stock price and the y axis shows the frequency.

We see a decline in frequency as the price increases. There is a large amount of stock with a low price. As there is an exponential distribution seen here, we look at log variables to perform the analysis.

Figure 2: binscatter plot of robinhood stock holdings vs stock market capitalization

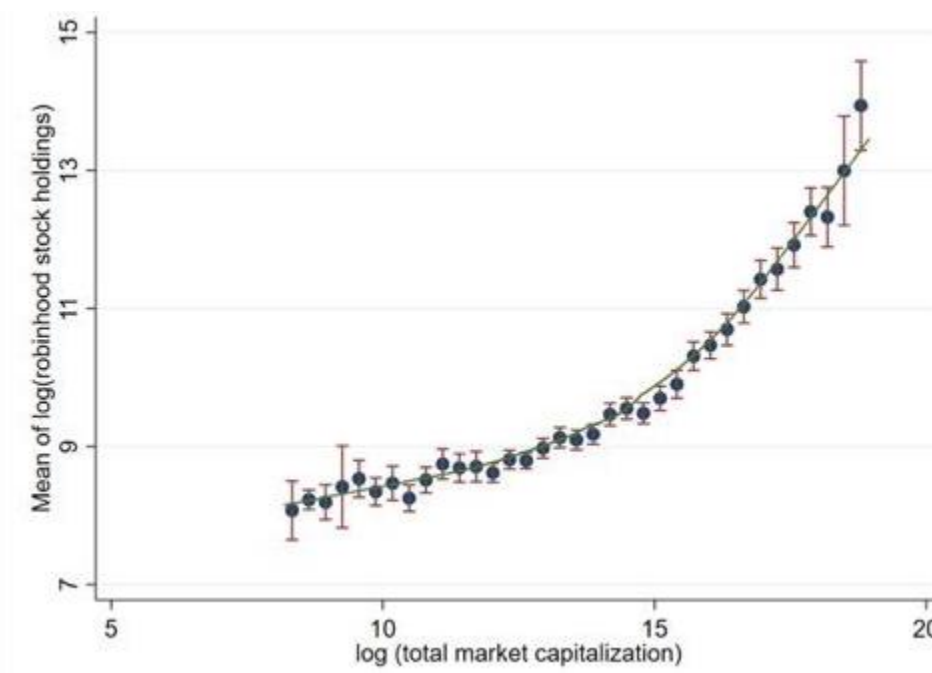


Figure 2 is a bin scatter of the log scale holdings of stocks in Robinhood vs the log scale total market capital of the stocks. Each bin contains x number of stocks with similar market capitalization, and the average is displayed as a point. 95% confidence interval calculated based on the variance of the holdings within each bin. The green line is the best fitted line according to the points.

We see a positive correlation between the total market capital of a stock and the total holdings of the stock in Robinhood. The total market capital of a stock represents in general how much the US stock market holds that stock. So it's reasonable that Robinhood users are holding corresponding more of that stock when a stock's total market capital is high. This is consistent with Robinhood investors behaving similar to a typical US investor.

The confidence interval is becoming larger as we look at bins with larger total market capitalization. That is because the number of observations is smaller within each bin as we look at larger stocks. This doesn't alter the clear pattern that high total market capital corresponds to higher total holdings.

Now that we've looked at how Robinhood investors behave similarly to representative US investors. We go on to look at the aspects in which Robinhood investors are different than US investors.

Figure 3: binscatter plot of robinhood stock holdings vs discussion on Reddit

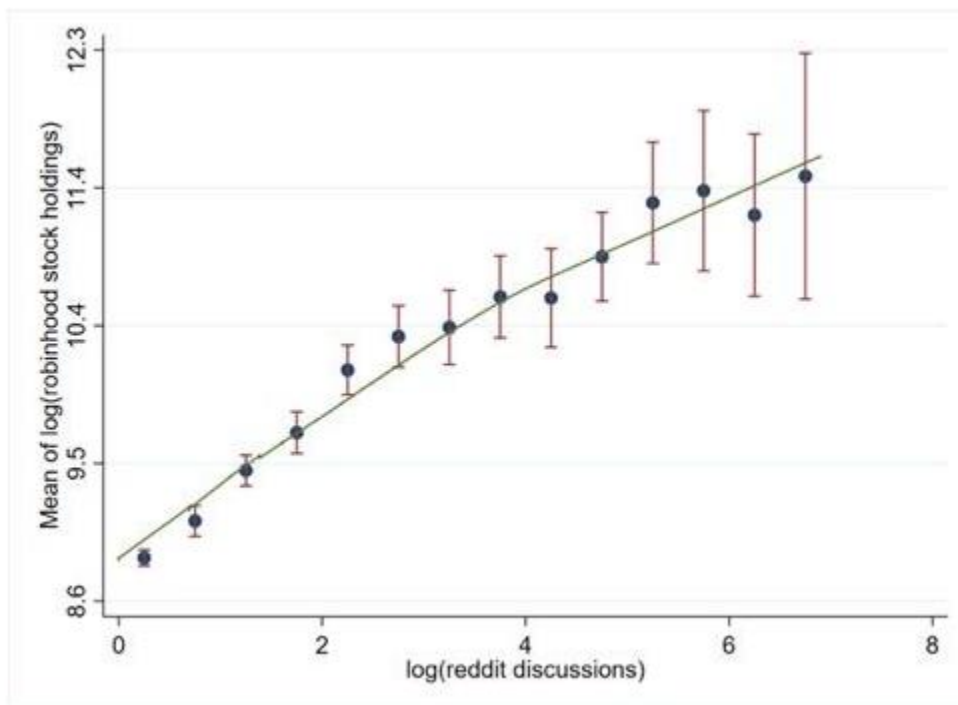


Figure 3 is a bin scatter of the log scale holdings of stocks in Robinhood vs. the log scale discussions on Reddit. Each bin contains x number of stocks with similar market capitalization, and the average is displayed as a point. 95% confidence interval calculated based on the variance of the holdings within each bin. The green line is the best fitted line according to the points.

We see a positive correlation between the total holdings of the stock in Robinhood and discussions on Reddit. For every 1% increase in log Reddit discussion the log of robinhood stock holdings also increases by roughly 1%.

Figure 4: binscatter plot of robinhood stock holdings vs discussion on Reddit, controlling for stock's market capitalization

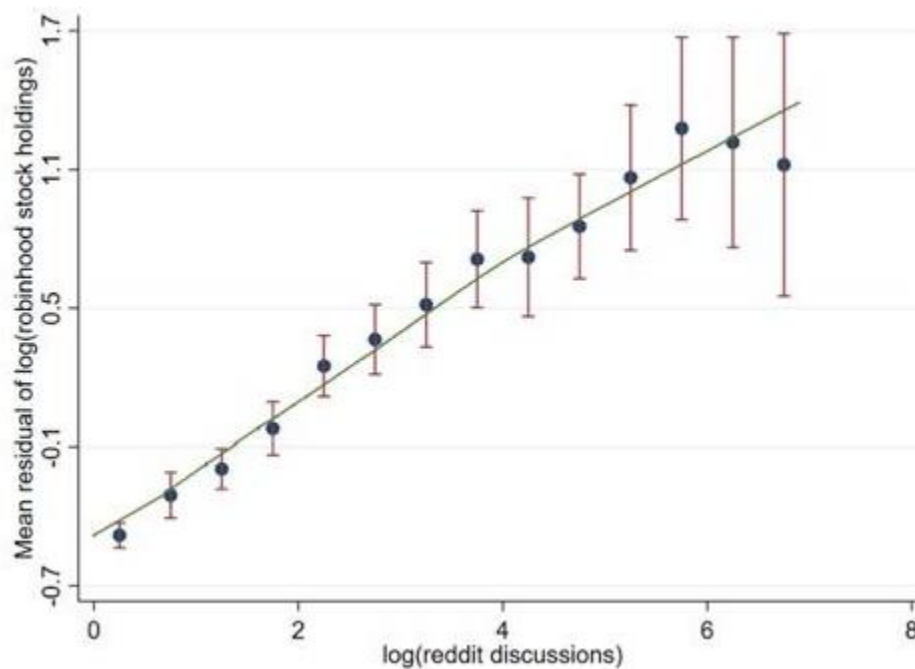


Figure 4 is the main result of this paper. We see from figure 2 and figure 3 that both market capitalization and Reddit discussion have a positive correlation with robinhood holdings of stock.

In figure 2 we learn that market capitalization has a strong correlation with robinhood user's holdings of a stock. It is reasonable to think that a stock that is high in market capitalization is also widely discussed on Reddit. To rule out the possibility that the positive relationship between Reddit discussion and robinhood investors holdings is driven by market capitalization, we show a bin scatter plot of the relationship between Reddit discussion and robinhood investors holdings while controlling for market capitalization.

After controlling for market capitalization, we still see a strong positive relationship between market capitalization and frequency of discussions of stocks onReddit. Based on this, we conclude that market capitalization plays an important factor here but still is not the full cause of the relationship. Even after controlling for market capitalization, discussion on Reddit is still a

strong predictor of stock holdings in robinhood.

In the classic model of rational investors, investors choose optimally which stock to invest in based on their risk tolerance and return expectation. Whether a stock is drawing people's attention on Reddit should have no impact on whether investors actually invest in the stock. However, figure 3 and 4 reveals that there is a lot of stock buying behavior when the stock is getting a lot of attention and discussion on Reddit, this is a violation of the view that retail investors are rational in their investment decisions. It is worth mentioning that in our analysis both positive and negative discussion of a stock would translate into a higher amount of Reddit discussions. Future research could study whether positive discussions of a stock results in more buy from investors than negative discussions of a stock.

IMPLICATIONS

As this paper establishes a positive relationship between frequency of discussions of a particular stock on Reddit and its market capitalization holdings, we learn that retail investors buy stocks that are frequently discussed on Reddit. This information can be useful for companies that rely on retail investors in their funding. They can mark some effort in advertising on Reddit in order to increase the number of retail investors buying their stock.

This paper also establishes that retail investors do respond to uninformative content on the internet which points to retail investor's irrational behavior. It is not ideal to see that retail investors just buy whatever stocks that grab their attention. If attention-grabbing stocks turn out to be a suboptimal choice for retail investors, this would lead to social inefficiency.

Attention-driven buying results from the difficulty that investors have searching the thousands of stocks they can potentially buy. The results of this paper could be beneficial for policy makers in financial regulations. Being aware of the fact that retail investors like to buy attention grabbing stocks, the government could come up with interventions to remedy this behavior. Many policies could be implemented to direct retail investors towards better information about their buying and selling of stocks. However, in order to have more effective policies, more research should be done in this field to know what causes attention-driven buying of stocks. By understanding the causes behind attention-driven buying, the Government will be able to create and implement effective policies to help mitigate the cost of bad investment decisions.

CONCLUSION

This paper contributes by providing the reader with a better understanding of the channels in which investors make investment decisions. The paper suggests that retail investors are

influenced by the discussions of stocks on internet forums. Our paper studies this by using stock discussion on subReddit r/wallstreetbet to proxy for information that retail investors receive from the internet. A positive relationship between the market capital holdings of Robinhood investors and amount of discussion around the stock on Reddit is found. This suggests that Robinhood investors tend to buy stocks with a high number of discussions on Reddit. This remains the same even after controlling for market capitalization of each stock. Based on this we are able to conclude that retail investors stock purchases are influenced by discussions on internet forums. This paper sheds light on the irrationality in retail investor's behavior and suggests the importance of intervention to correct it.

REFERENCES

Antweiler, Werner, and Murray Z. Frank. "Is all that talk just noise? The information content of internet stock message boards." *The Journal of finance* 59.3 (2004): 1259-1294.

Lavrenko, Victor, et al. "Mining of concurrent text and time series." *KDD-2000 Workshop on Text Mining*. Vol. 2000. Citeseer, 2000.

Tumarkin, Robert, and Robert F. Whitelaw. "News or noise? Internet postings and stock prices." *Financial Analysts Journal* 57.3 (2001): 41-51.

Das, Sanjiv R., and Mike Y. Chen. "Yahoo! for Amazon: Sentiment extraction from small talk on the web." *Management science* 53.9 (2007): 1375-1388.

Riffin, Nor Azua Mhd, and Noryati Ahmad. "A conceptual paper on factors affecting investment decisions of Malaysian investors." *2012 International Conference on Innovation Management and Technology Research*. IEEE, 2012.

Uslu Divanoğlu, Sevilay, and Dr BAĞCI. "Determining the Factors Affecting Individual Investors' Behaviours." *International Journal of Organizational Leadership* 7 (2018): 284-299.

Peress, Joel. "The media and the diffusion of information in financial markets: Evidence from newspaper strikes." *The Journal of Finance* 69.5 (2014): 2007-2043.

Wysocki, Peter D. "Cheap talk on the web: The determinants of postings on stock message boards." *University of Michigan Business School Working Paper* 98025 (1998).