Essays in Empirical Asset Pricing

by

Chaehyun Pyun

(Under the Direction of Jeffry Netter)

Abstract

This dissertation includes two chapters on empirical asset pricing. In the first chapter, I study how data from different social media platforms may have different market implications. Using data from investing-related chat rooms, I find that social media group investing can help investors find high alpha stocks and are more informative than individual posts and comments in asynchronous investing forums. The second chapter examines various aspects of the sensitivity of the risk-neutral excess stock variance attempting to explain expected returns to reasonable alterations in the empirical design. The results reveal substantial time-series and cross-sectional variations in the predictive relation between risk-neutral variance and future stock returns.

Keywords: Empirical Asset Pricing, Financial Markets, Equity Returns

Essays in Empirical Asset Pricing

by

Chaehyun Pyun

B.S., Sogang University, South Korea, 2011

M.A., Yonsei University, South Korea, 2014

M.A., University of Illinois at Urbana-Champaign, 2015

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2022

Essays in Empirical Asset Pricing

by

Chaehyun Pyun

Major Professors: Jeffry Netter

Committee: Zhongjin Lu
Annette Poulsen
Jie He

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2022

To my wife, Jiyoon Lee, who stood by me through the lows and the highs of graduate school, our beautiful daughter Zoey Haim Pyun, and to my parents, Sanghoon Pyun and Sunghee Kang, who raised me into who I am. Also, to Seung-hyun Pyun, Bok-Yang Pyun, the Ha Family, the Lee Family, and Geum-hi Suh for their moral support and help along the way.

Acknowledgements

I express my deepest appreciation to my dissertation committee members Professor Jeffrey Netter, Annette Poulsen, Zhongjin Lu, and Jie He, for their wisdom, direction, support, and encouragement. Special thanks go to Zhongjin Lu, that made this dissertation possible.

I would also like to thank other faculty members, including Paula Suh, Steven Malliaris, Kayla Freeman, Pekka Honkanen, and Filipe Correia, for their insightful suggestions, as well as seminar participants Wayne Ferson, Junbo Wang, Ryan Israelsen, and Alex Chinco for discussing my paper and providing valuable comments.

Last but not least, it was a great honor and pleasure to join this journey with my cohort members, William Bost, Bo Sun, and Cosmin Cioanta, in the pursuit of knowledge.

Table of Contents

Chapter 1

Social Media Group Investing

1.1 Introduction

Self-directed investment clubs make investing and trading sociable and have the advantage of having members with knowledge and experience in different sectors, allowing collective investment discussions.[1] Does collective brainpower from live group discussions improve security selection ability? However, not much research has been done regarding investment clubs due to the difficulty of collecting recordings of offline investment discussions.[2]

As newer generations become more computer-savvy and actively use social media, investors are searching for more intimate, community-focused environments to exchange trading ideas and share their latest trades and strategies. After COVID-19, online group discussions from these "social media-savvy" investors significantly increased because of travel restrictions and stay-at-home orders. This explains the significant rise in investors using Discord, a real-time communication platform that allows live interaction between groups of investors and has many similarities with self-directed investment clubs.[3] Each Discord investing server has its own

[1] The U.S. Securities and Exchange Commission defines a "self-directed investment club" as a group of people where members research and select investments together but invest individually instead of pooling their money.

[2] Barber and Odean (2000) find that 60% of investment clubs underperformed the market index, but they only
[2] Barber and Odean (2000) find that 60% of investment clubs underperformed the market index, but they only
examine a randomly selected small sample of 166 investment clubs, despite the fact that the National Association of Investors Corporation (NAIC) reported over 35,000 active investment clubs at the time.

[3] See "Retail investors are flocking to Discord amid the Reddit day trader revolution," Market Business Insider News. February 13, 2021.

unique name, and investors can join the server that best fits their financial objectives. With U.S. adults increasingly spending more time on Discord,[4] it may arguably be becoming one of the most popular platforms with group chatting capability for investing.[5]

Rather than looking into self-directed investment clubs, the benefits of investigating Discord investing-related servers are twofold. First, it is possible to obtain a comprehensive data set while exploring the collective wisdom of live group discussions. Second, the live chat feature and moderators within each investing server help reduce noise from spam and bots. As a result, the data collected from Discord becomes a much more accurate data set.

I call the process of investors gathering in groups in social media and discussing investment opportunities together via live group chat "social media group investing," a concept similar to the online version of self-directed investment clubs. Previous studies have mainly focused on posts and comments by individuals. Thus, traditional platforms on which investors browse through old posts to find relevant information such as Reddit WallStreetBets (WSB hereafter)[6] are not consistent with this definition of social media group investing. Some of the investing-related servers in Discord even accept applications to join their waitlist, which is another aspect that differentiates itself from other popular social media platforms.[7]

---

[4] As of May 2021, there are more than 21,500 investing servers at Discord. A list is provided here. https://stockbot.us/pages/servers.php

[5] See "What Is Everybody Doing on Discord?" The Wall Street Journal. March 8, 2021.

[6] WallStreetBets is a community on Reddit that features daily posts outlining which stocks and options participants plan to invest in the following day.

[7] See (https://www.hashtaginvesting.com/blog/best-discord-servers-of-group-chats-for-stock-trading-investors) for some examples.

The first part of my paper empirically examines the impact of social media group investing on trading during a period of market uncertainty. By observing which stocks are the most heavily discussed from investment-related servers each day, I create a relatively direct measure for popular stocks. The unique data set in this paper reveals that "popular stocks" with lower mentions have more variety, stay popular for a shorter period of time, and are smaller in size.[8] I find that popularity has a positive relationship with future stock volatility, abnormal volume, and returns. I separately examine subsamples of below-median and above-median firm size for each day and find that the impact of popularity is greater for smaller firms, consistent with social media-savvy investors reducing the informational asymmetries associated with less visible stocks. I also find that higher returns are driven by the continuity of remaining popular than by the day the stock becomes popular.[9]

During the first part of my paper, I also investigate two stories of channels in why more discussions may be associated with increases in stock price. One possibility is that investors understand what will happen to the cash flows of those most-mentioned companies by doing some analysis, which is related to the informed trading explanation. The other possibility is that investors are pumping the stock, which is related to the price pressure explanation. In other words, is the popularity deserved or undeserved?

To answer this question, my paper examines buy-and-hold returns over longer horizons and looks into insiders' incentives and opportunities. Empirical results show that the buy-and-

---

[8] "Popular stocks" refers to specific stocks investors currently mention and discuss the most on social media platforms.

[9] This helps mitigate concerns that material news and events drive the results. Section 1.3 provides more details addressing endogeneity concerns.

hold returns display a steady, continued upward drift over long horizons, providing empirical support for the informed trading explanation. Moreover, I find that insiders are less likely to sell shares after a stock becomes popular, lending further support for informed discussions from a platform with live group chatting capability.

The second part of my paper investigates whether social media group investing through live group chat provides better returns. In other words, is stock selection from a synchronous platform better than searching and scrolling from an asynchronous platform? My paper is the first to examine Discord and create a portfolio that can be contemporaneously traded using the basket of stocks most heavily discussed in live group chat rooms. Using equal-weighted (EW hereafter) and value-weighted (VW hereafter) portfolios, I show how investing in these portfolios would have progressed over time. I choose Reddit WSB as the asynchronous platform since I find that it shows a similar rise in users and the overall amount of chatter. The cumulative portfolio returns show that the Discord EW portfolio outperforms the S&P 500 and both Reddit WSB EW and VW portfolios, regardless of the portfolio rebalancing frequency. Results further show that small and mid-cap stocks perform well in the Discord portfolio. Sentiment analysis further shows that the sentiment for heavily discussed stocks is mostly positive, making the tone of the stock discussions less of an issue than identifying "which" stocks investors are talking about the most.

I also find that "meme" stocks like GameStop and AMC are mentioned more in Reddit than in Discord.[10] These meme stocks also become one of the most discussed stocks several months earlier in Reddit. For example, even though the overall chatter on both platforms is similar in magnitude, GameStop was not a hot topic in Discord until the January 2021 short squeeze, while it was one of the most talked about stocks on Reddit since October 2020.

With respect to the Fama-French benchmark models, I find that the Discord EW portfolio did not underperform, with alphas larger on nearly all models compared to other portfolios and tilts towards small growth stocks. Additionally, I find that Reddit WSB-popular stocks have nearly three times higher loading on Bitcoin returns than Discord-popular stocks. This result is consistent with Reddit WSB-popular stocks containing many meme stocks, suggesting that the Reddit WSB popularity is possibly more correlated with undeserved popularity.[11]

In short, the second part of my paper demonstrates that it is beneficial for investors to join live group chats. Comparing two trending social media platforms also answers the question "Where is the smart chatter?" and shows that selecting the right platform can meet the needs of both retail and institutional investors. Results suggest that social media that allow live interaction between groups of investors like Discord may be better for investors who want to produce higher alphas. On the other hand, the Reddit WSB portfolio underperforms and has much more mentions of meme stocks on the platform. This suggests that Reddit WSB does not discuss as many good trade plays and is better suited for institutional investors looking for a leading

[10] I categorize a "meme" stock as referenced by the stock market community as a stock with heavy short interest that can be artificially manipulated.
[11] See "Bitcoin's wild rally—and a fear of missing out—has retail investors flocking to crypto," CNBC. January 8, 2021.

indicator for the next GameStop frenzy so that no stocks in which they are short are heavily promoted.[12] When it comes to popular stocks, social media-savvy investors tend to favor small-cap companies with high growth potential, indicating that Discord's social media-savvy investors are ready to put in the effort to investigate and uncover promising small-cap companies that have less focus and attention than large-cap companies. The live group chat functionality with moderators, faster information with Bayesian updating in beliefs, and the users' incentives may contribute to explaining Discord's smarter chatter with better stock selections.

My paper contributes to two lines of research. First, it adds to the literature analyzing the effects of social media on the stock market. Several studies use Seeking Alpha articles, StockTwits messages, and Twitter tweets and find a positive relationship with earnings or returns (Chen, De, Hu, and Hwang, 2014; Renault, 2017; Bartov, Faurel, and Mohanram, 2018; Farrell, Green, Jame, and Markov, 2021), while others do not find such relationships (Heimer, 2016; Giannini, Irvine, and Shu, 2018; Cookson, Engelberg, and Mullins, 2020). Previous studies mainly focused on social media with posts and comments by individuals. However, many social media-related studies do not account for the fact that many messages are posted by bots and can be spam. For example, researchers at Carnegie Mellon University discovered that nearly half of the Twitter accounts spreading messages about COVID-19 are most likely bots.[13] My paper fills a gap in the literature by introducing a social media platform where bots and spam are less of a concern and allows real-time group discussions to address new questions. I use a novel approach

[12] In January 2021, armies of retail investors on Reddit WSB were hunting the so-called wolves of Wall Street with massive GameStop buying campaigns. This drove up the price of shares and call options, forcing short sellers to repurchase GameStop shares sooner to avoid more losses as the stock rose.
[13] See https://www.cmu.edu/news/stories/archives/2020/may/twitter-bot-campaign.html

to study social media popularity and provide insight into how different platforms may have different market implications during a period of high market uncertainty.

Second, my paper also contributes to the literature on retail trading and investing skill. Several studies either found or failed to discover evidence of retail trading's informativeness. (Hvidkjaer, 2008; Kaniel, Liu, Saar, and Titman, 2012; Farrell, et al., 2021; Eaton, et al., 2021; Boehmer, et al., 2021). While earlier work in the related literature uses a broader measure of retail investors, this paper studies social media-savvy retail investors. Moreover, rather than trying to make a general conclusion that social media-savvy investors are either informed or uninformed, my paper focuses more on where the smart chatter is and whether the discussions from that platform are informed. This paper shows that Discord is better suited to social media-savvy investors seeking high returns, while Reddit WSB is better suited to institutional investors looking to protect their portfolios. More importantly, the upward drift from Discord-popular stocks over long horizons and results from insiders' trades provide more support for the informed trading explanation than the price pressure explanation.

1.2 Data and Descriptive Statistics

1.2.1 Background Information

Discord is a communication software that allows people to live chat 24/7 via text, voice, or video. It is designed for real-time group communication, not for scrolling through old posts. On other platforms, people have to post an article or upload what they want to say in their account, which is not live interaction among people. Also, Discord does not have news feeds and does not sell advertisements, whereas other platforms are mostly based on advertising business models.

Investors are increasingly using Discord to swap trading tips and ideas. This makes Discord a rich space to investigate up-to-date findings regarding social media-savvy investors. The simultaneous group chatting feature of Discord overcomes the weakness of Seeking Alpha in connecting all investors to each other and may be a reason why Seeking Alpha is more suitable for influencers to share their insight to their own audiences. Discord investing servers have many similarities with self-directed investment clubs and enable looking into social media group investing. In 2020 and beyond, Discord servers have become one of the best places to socialize and discuss trades.

When users create an account in Discord, they can join a particular investment-related server either by receiving an invitation link from a friend who is already in that server or by browsing through a list of popular servers. Each investing server has its own unique server name, and some of them even have a waitlist to join. When users join a server, they must prove that they are a real person and agree to the server rules regarding content and behavior. To maintain an educational and informational community, moderators monitor the group chat and dismiss users that they feel are unhealthy to the community.

Live interaction among people and moderators makes it easier for Discord servers to detect and prevent spam, bots, and fake accounts. Therefore, it is more likely that Discord is less contaminated with noise or other unwanted features in the data than other social media platforms. Other platforms may need additional data cleaning processes. Also, when the objective is to determine which stocks people are most actively talking about in a given timeframe, the better choice is to choose the platform with live group chat rooms where investors bounce ideas off of other traders in real-time and continuously discuss their investment plans.

1.2.2 Main Variable Construction and Data Sources

Previously, studying the impact of popularity was difficult because the only direct way to gather information on which stocks are popular at any given time would be to gather thousands of investors and ask each of them which stocks are on their watchlist on a daily basis. My popularity measure overcomes the limitations of most indirect proxies by similarly following and accomplishing an approach mentioned in Barber and Odean (2008) that was deemed impractical at the time:

> How can we measure the extent to which a stock grabs investors' attention? A direct measure would be to go back in time and, each day, question the hundreds of thousands of investors in our data sets as to which stocks they thought about that day. Since we cannot measure the daily attention paid to stocks directly, we do so indirectly. (Barber and Odean, 2008)

Social media has made data much more observable to researchers. By observing which stocks investors are discussing the most in real-time investing group chat rooms every day, I create a direct measure for popularity that facilitates research on its implications. I call the top 50 most discussed stocks from group chats per day "popular" stocks.[14] The intuition here is that as a stock's popularity and retail interest grow, investors talk about it more. A notable aspect is that I collect it from retail chatter, not from commonly used data providers that need subscriptions. Not only does this make the data more interesting, but this feature also provides the advantage of allowing this study to be easily updated in the future with less concern about data availability.

To see any difference in the amount of chatter about stocks during different periods of the day, I collect stock mentions using an application programming interface (API hereafter) for three different horizons. The first horizon is the 24-hour period from market open on day $t-1$ to market open on day t (OTO). I then divide this into two parts: market close on day $t-1$ to market open on day t (CTO) and market open on day t to market close on day t (OTC). To eliminate duplicate counts, any particular symbol is recorded once per day for each Discord member, no matter where or how many times a member may use it during the day. To compare with a social media platform that does not have real-time group chat rooms, I also collect mentioned symbols from Reddit WSB.[15]

Figure 1.1 shows the total count of the top 50 stocks most-mentioned per day, from investing group chat rooms in Discord and the Daily Discussion thread of WSB from Reddit. Panel A looks at the difference between the two platforms, and Panel B looks at the difference between the collection period on a day. Panel A shows that the number of investors communicating investment opportunities through Discord and Reddit increases over time. Compared to Reddit WSB, the trend line and plotted points from Discord display a monotonic

[14] For most days, stocks that have not been talked about much come after this cutoff.
[15] The WSB tickers are collected from an API provided by Quiver Quantitative.

increase in user mentions over time. An exponential rise in counts is observed in January 2021, when the GME short squeeze occurred. Panel B shows that investors discuss stocks more during market hours (OTC) than when the market is closed (CTO), as illustrated by the steeper slope of the trend line. It is interesting also to see that stock discussions did not stall during the COVID-19 crisis. This suggests that investors did not lose interest in investing from the sharp market-wide downturn, which is consistent with the increase in aggregate Robinhood holdings throughout the COVID-19 crisis in Ivo Welch (2021).

Other data sources used in my paper are Compustat for firm characteristics, Capital IQ for material news and events, Quiver Quantitative for historical sentiment data, Thomson-Refinitiv for insider activity, and Ken French's website for the risk-free rate and factors.[16] I exclude microcap stocks, that is, those with a price below $5 a share or a market capitalization below $300 million from the sample to mitigate microstructure issues.

The sample period starts from January 2020 to January 2021. Since this study is more focused on looking at the stock market following COVID-19 when there was a massive increase in retail investors and Discord and Reddit gained greater popularity, the sample period satisfies the scope of this study. However, as cautioned in Ivo Welch (2021), each reader must make their own subjective judgment of which results are likely to have external validity, as in any empirical study. Nevertheless, a virtue of investigating the unique year of 2020 is that it allows readers to have a better picture of how investors react and what they discuss during stressed times.[17]

---

[16] The data is hand-collected for missing dates in the Capital IQ database.
[17] For future research, it would be interesting to compare with less volatile times (e.g., post-COVID-19 era).

The OTO returns used in my paper are calculated by,

$$r_{OTO,t} = \frac{P_{open.t+1}}{P_{open,t}} - 1, \tag{1}$$

where $P_{open,t}$ is the first trade price on day t on any of the exchanges on which the security is traded.

1.2.3 Monotonic Relationships

This subsection investigates the relationship between the rankings of the most-mentioned stocks with the number of unique tickers that appear, the number of days they remain in the most-mentioned list, and firm size. I sort popular stocks based on their mention rankings into five deciles (the highest tier of 1–10 to the lowest tier of 41–50) for each day and aggregate them across the sample period.

Table 1.1 shows several monotonic relationships that hold in both Panels A and B. As the rank tier goes up from low tier (41–50) to high tier (1–10), there is an inverse relationship with the number of unique tickers, that is, the same popular stocks over the sample period are repeatedly placed in the same tier. It also shows a positive relationship with days included in the most-mentioned list, which implies that stocks that are popular tend to remain in that state longer. Lastly, a positive relationship is observed with firm size, suggesting that stocks in the higher tiers are, on average, larger in size. In sum, popular stocks with higher mentions tend to be repeatedly placed in the high-mentioned rankings, are larger in size, and stay popular longer, while popular stocks with lower mentions have more variety, are smaller in size, and stay popular for a shorter period of time.

1.3 Impact of Social Media Group Investing on Trading

The first part of my paper empirically examines the impact of social media group investing on trading during a period of high market uncertainty. Shares in companies popular on social media may or may not deserve their rising market capitalization. The latter should be true about meme stocks, for which the price of a stock soars to a level that does not reflect the reality of the company. Therefore, while social media group investing can provide investment value, investors can still be exposed to undeserved popularity. I try to distinguish undeserved from deserved popularity as part of my second analysis. One challenging task is ensuring that the market is reacting more to the popularity around the stock than to news articles covering a corporate event about the firm. I address this concern and make an adjustment to the main analysis to better isolate the effect of popularity.

To begin with, according to Ivo Welch (2021), Robinhood investors overweighted some rather unusual portfolio positions, supporting the assumption that most of the popular stocks may not be related to corporate events. For example, Ivo Welch (2021) mentions that cannabis stocks were remarkably popular among Robinhood Investors. At the end of January 2019, Aurora Cannabis (ACB) was briefly the most widely held stock, with Apple (AAPL) second at the time.

Moreover, less than one-third of the popular stocks from Discord and Reddit WSB are S&P 500 companies, suggesting that news article dissemination is likely an issue for most of the popular stocks in my paper. The reason is that firms that are not included in the market index are less likely to be mentioned in news articles, supporting the view that results should not be driven by media focus.

Most importantly, one of the main findings in my paper is that higher returns are driven more by when a stock continues to stay popular than when it becomes popular. It would be more reasonable to expect the opposite outcome if material news and events drive the results. Moreover, the returns do not show statistical significance immediately the day after a stock becomes popular. Nevertheless, I use the Key Developments data from Capital IQ to search for any material news and events that may affect the market value of a stock each day and see if any of these events coincides with the day a specific stock receives the most mentions. I drop these occurrences from the sample to help mitigate concerns about confounding variables explaining the main results of my paper. I also control for several past performance variables commonly used in the literature in the main analysis.

While different recording time frames are used in previous subsections, the CTO period is selected to collect the most heavily discussed stocks in live group chat rooms to create the "popularity" variable. The reason is that firms tend to submit important regulatory filings and most earnings announcements occur outside of the regular market open times. Therefore, collecting the most-mentioned stocks during this period improves accuracy since this is when investors have all the available information needed to logically plan out their investment decisions for the next trading day. Also, it is most likely the best time for investors to look back and analyze what has happened during market open hours. Therefore, popular stocks are collected from market close on day $t-1$ to market open on day t, and OTO portfolio returns are computed from market open on day t to day $t+1$.

1.3.1 Trading Volume and Stock Volatility

To test whether popularity predicts trading volume and stock volatility, I use Fama and MacBeth (1973) (FMB hereafter) regression used in asset pricing models and is compatible with many assets across time. Throughout the main analyses, I drop popular stocks that had any material news and events that may affect the market value of the stock for each day to mitigate endogeneity concerns.

The regression is given as

$$DepVar_{i,t} = \alpha_i + \beta_{1,i} POPULAR_{i,t-1} + \sum_{k=1}^{K} \beta_{2,k,i}\, Control_{k,i,t} + \varepsilon_{i,t}, \tag{2}$$

where $\mathrm{DepVar}_{i,t}$ is abnormal trading volume (AbVolume), measured by firm *i*'s log trading volume on day *t* minus its average log trading volume on days *t*−5 to *t*−1 and stock volatility (|AbRet|), measured as the absolute value of firm *i*'s abnormal stock return on day *t*. $\mathrm{POPULAR}_{i,t-1}$ is an indicator variable which equals one if the stock is one of the top 50 most-mentioned stocks on Discord from market close on day *t*−1 to market open on day *t* (CTO). $\mathrm{Control}_{k,i,t}$ are log of the trading volume on day *t*−1, VOLUME(*t*−1), firm *i*'s log of market capitalization on day *t*−1, $\mathrm{SIZE}_{i,t-1}$, cumulative returns from day *t*−5 through day *t*−1, $\mathrm{Ret}_{i,t}[-5,-1]$, and the log of Amihud's (2002) illiquidity measure averaged over days *t*−5 to *t*−1, $\mathrm{ILLIQ}_{i,t}[-5,-1]$. The daily illiquidity measure is equal to $10^6 * |\mathrm{Ret}_{i,t}|/\mathrm{DVol}_{i,t}$, where $\mathrm{DVol}_{i,t}$ is the stock's dollar volume. These control variables are commonly used for stock characteristics in the literature (Tetlock 2011; and Sprenger, Tumasjan, Sandner, and Welpe 2014). Other control variables include BOOK-TO-MARKET, computed as the log of the ratio of book equity from the most recent fiscal year to the market capitalization at the end of the year, PROFITABILITY, computed as operating income before depreciation as a fraction of average total assets based on most recent two periods, past VOLATILITY, computed as the log of the standard deviation of daily returns from the previous month, and MOMENTUM, measured by the return over the past 12 months without the most recent month's return (Jegadeesh and Titman 1993; and Martin and Wagner 2019).

Table 1.2 presents results from daily FMB regressions of abnormal volume (AbVolume) and stock volatility (|AbRet|) on popularity and other control variables, shown in equation (2). I also conduct subsample analysis by separately examining subsamples of below-median and above-median firm size for each day.

The main finding in Table 1.2 is that popularity has a positive and significant effect on abnormal volume and stock volatility on the next day for all specifications. In particular, the subsample analysis shows that the effect on abnormal volume is stronger in small firms (0.485) compared to big firms (0.196). The coefficient on small firms translates to popular small firms having about 60% more trade volume on average on the following day compared to non-popular small stocks. Similarly, this corresponds to popular big firms having about 20% more trade volume on average on the next day compared to non-popular big stocks. The difference between the subsamples is more apparent when the dependent variable is stock volatility. The coefficient is five times as strong in small firms (0.035) than in big firms (0.007).

1.3.2 Deserved and Undeserved Popularity

This subsection attempts to identify the difference between deserved popularity and undeserved popularity. Undeserved popularity is related to the upward price pressure explanation that suggests positive future stock returns are due to uninformed demand shocks. In that case, emotions can drive prices up, and a reversal can be observed as the stock price eventually returns to prices that are in line with economic fundamentals. Conversely, deserved popularity is related to the informed trading explanation, where instead of a reversal, an upward drift is observed. I investigate these explanations by looking into buy-and-hold returns of popular stocks over different holding periods. This analysis also allows my paper to explore how long investors should hold popular stocks to maximize returns. The FMB regression for predicting returns for different holding periods is given as

$$Ret[a,b]_{i,t} = \alpha_i + \beta_{1,i} POPULAR_{i,t-1} + \sum_{k=1}^{K} \beta_{2,k,i}\, Control_{k,i,t} + \varepsilon_{i,t}, \tag{3}$$

where $\mathrm{Ret}[a,b]_{i,t}$ is the OTO return in equation (1) compounded from day t+a to t+b. This regression uses Newey and West (1987) standard errors adjusted up to two times the holding period to account for overlapping OTO return observations.

Table 1.3 shows that popular stocks get greater returns compared to stocks that are not popular. Interestingly, the *t*-statistic increases to 2.06 when investors buy popular stocks two days after they become popular, compared to a *t*-statistic of 1.63 when buying the day after. The statistical significance slightly increases for longer holding periods. The steady increase in returns over longer horizons suggests a significant upward drift. For example, holding for 30 days increases the effect from 0.3% to 4%. When holding for more than two months, the buy-and-hold return for popular stocks is 6.6% higher than those not popular. Overall, the steady and continued upward drift over long horizons provides support for the informed trading explanation.

The buy-and-hold results suggest that market participants who buy popular stocks and hold on to their investments pay off. However, caution is still advised because the return performance of any investment strategy, whether or not it is published in prestigious finance journals, is a notoriously poor predictor of future investment performance in general (Welch, 2021). Nonetheless, it is still interesting to see from Table 1.3 that there is no obvious optimal holding period. Therefore, it may be a good idea for investors to connect to social media and find out which stocks everyone is talking about, regardless of their investment horizon.

1.3.3 Entry and Continuity of Popularity

Section 1.3.2 shows an upward drift in future stock returns. To better examine whether the statistical significance comes from stocks that just became popular on that day or stocks that are already popular, for each stock, I record the day a stock suddenly appears in the top 50 most-mentioned list and also separately record the consecutive days it remains in the list. I then run the following FMB regression:

$$\begin{aligned} Ret[a,b]_{i,t} = \alpha_i + \beta_{1,i} ENTRY_{i,t-1} + \beta_{2,i} CONTINUITY_{i,t-1} \\ + \sum_{k=1}^{K} \beta_{3,k,i}\, Control_{k,i,t} + \varepsilon_{i,t}, \end{aligned} \tag{4}$$

where $\mathrm{Ret}[a,b]_{i,t}$ is the OTO return in equation (1) compounded from day *t*+*a* to *t*+*b*. $\mathrm{ENTRY}_{i,t-1}$ is an indicator variable that equals one if the stock was not in the Discord most-mentioned list on day t−2 and entered the Discord most-mentioned list on day t−1. $\mathrm{CONTINUITY}_{i,t-1}$ is an indicator variable that equals one if the stock was in the Discord most-mentioned list on day t−2 and also on day t−1.[18] Newey and West (1987) standard errors adjusted up to two times the holding period are used to account for overlapping OTO return observations.

Table 1.4 shows that the higher returns do not come from "fresh" popular stocks, that is, stocks that become popular which previously were not. Rather, it comes from stocks continuing to stay popular. The estimates and t-statistics of $\mathrm{CONTINUITY}_{i,t-1}$ are larger than $\mathrm{ENTRY}_{i,t-1}$ for all columns. These results mitigate the concern that specific events on a certain day are driving the results. The results imply that staying popular is a more important driver for higher returns.

1.3.4 Insiders' Incentives and Opportunities

Section 1.3.2 shows that between deserved and undeserved popularity, the most discussed stocks from Discord chat rooms are closer to deserved popularity, with informed trading as the more likely explanation. This subsection looks into insider trades to see whether their trades further support the informed trading explanation.

Many researchers have found that corporate insiders' trades are informative and predict future abnormal returns (Datta and Datta, 1996; Lakonishok and Lee, 2001; Biggerstaff, Cicero, and Wintoki, 2020). The intuition behind the signals from insider trading is that insider purchases and sales are motivated by an informational advantage (Biggerstaff et al., 2020). Since

[18] In untabulated results, regressions using the CONTINUITY variable as a continuous variable (fraction of previous week in the most-mentioned list) yields similar results.

Section 1.3.2 demonstrates that popularity benefits shareholders by increasing returns, if insiders sell more after a stock becomes popular, this weakens the informed trading explanation and can be attributed to managerial opportunism, namely the use of popularity for shares to be sold at higher prices. In contrast, if insiders sell less or buy more, it supports the informed trading explanation, suggesting that positive private information is slowly being embedded into prices.

Following Biggerstaff et al. (2020), I include officers, directors, and 10% beneficial owners of a company's stock as an "insider," which accords with the definition under S.E.C. regulations. I collect open market purchases and sales by insiders and utilize two variables to assess daily insider trading activity at each firm, following Lakonishok and Lee (2001) and Blackburne, Kepler, Quinn, and Taylor (2020). InsiderSell is an indicator variable equal to one if insiders at the firm are net sellers, that is, the number of shares sold exceeds purchases on that day and zero otherwise. The other variable is InsiderNSR, which is the ratio of net sales to total insider transactions. Using these variables as the dependent variable, I estimate the following pooled panel regression:

$$DepVar = \alpha + \beta_1 Day[-10,-1] + \beta_2 Day[0] + \beta_3 Day[+1,+10] + \sum_{k=1}^{K} \beta_{4,k}\, Control_k + \varepsilon, \tag{5}$$

where Day$[a,b]$ is an indicator variable that equals one if the day falls within the window and zero otherwise. Day[0] is an indicator variable that equals one if the stock was not in the Discord most-mentioned list the previous day and entered the list on day t=0. Similar to Blackburne et al. (2020), I estimate equation (5) with and without fixed effects and clustered standard errors. Since the window variables can overlap if a firm becomes popular again within a short period, I also retest equation (5) after removing overlapping window days.

Table 1.5 shows that across all specifications, there is no sign of insiders selling their shares after a stock becomes popular. When fixed effects are not included, and with non-overlapping windows, all window variables are negatively statistically significant for both dependent variables. The post-popular window is statistically significant for all specifications, even when fixed effects are included, with *t*-statistics of −2.47 and −2.50 when the dependent variable is InsiderSell and InsiderNSR, respectively. These results are inconsistent with managers exploiting popularity to sell their shares at higher prices. Results suggest that insiders are less likely to sell shares in a short window after a stock becomes popular, providing further support for the informed trading explanation.

1.4 Synchronous vs. Asynchronous Platform

The second part of the paper investigates whether social media-savvy investors discuss better stocks during live group chats in a synchronous platform and discuss possible explanations. I chose Reddit WSB as the asynchronous platform because it has a similar increase in users and the overall amount of chatter. Stocks are collected from Reddit WSB from the posts and comments of the discussion thread.[19] Prior to conducting any portfolio analysis, I evaluate how the stocks that are discussed differ between platforms. I then observe the cumulative OTO

[19] The 24-hour period of the previous day is used to collect ticker mentions for Reddit due to data constraints. Using different recording time frames for Discord gives similar results, which implies the data constraint for Reddit is neglectable. Moreover, Reddit is not the main focus of the study and is used for comparing purposes only.

portfolio return using equation (1) for the Discord portfolio using both the VW and EW schemes and see whether it outperforms the S&P 500 over time using different rebalancing frequencies.[20] I also check the performance with respect to the Fama-French benchmark models by observing the net of risk-free portfolio return performance. In addition, an OLS regression is used to test whether the popular stocks are correlated with the performance of Bitcoin. Finally, I compare the results using sentiment analysis.

1.4.1 Overall Most-mentioned Stocks

I see whether the most-mentioned stocks vary across platforms and recording timeframes and any noticeable industries represent the sample period. The percentage column in Table 1.6 represents the total number of times a stock ticker is mentioned divided by the total number of times the top 50 stocks are mentioned during the sample period.

Table 1.6 shows the 15 most-mentioned stocks over the sample period, and it is visible that many of them are similar across platforms. However, the majority of tickers that represent less than 1% differ greatly between platforms, which are not shown in Table 1.6. Panel A shows that meme stock tickers such as GME, AMC, BB, and NOK are mentioned more in Reddit than Discord. For instance, GME is the second most-mentioned ticker on Reddit during the sample period, while it is the tenth most-mentioned on Discord. This suggests that Robinhood investors gathered on Reddit and encouraged each other to pile into these meme stocks, which eventually intensified losses among professional traders. The Global Industry Classification Standard (GICS)

[20] Many investors and fund managers compare their performance against the S&P 500 index because it is the best-known market proxy for the U.S. stock market. See https://www.investopedia.com/terms/m/market-proxy.asp

shows that the technology, software, semiconductor, retail, and entertainment industries defined the sample period.

1.4.2 Most Mentions by Month

In this subsection, I collect which stocks have been mentioned the most for each month. This allows seeing if there are any differences in the sectors being focused on between Discord and Reddit at a given period of time. This also allows readers to see how the popularity in specific sectors has changed over time.

Table 1.7 shows that during the COVID-19 crash in March 2020, Discord focused on some pharmaceutical companies (Inovio Pharmaceuticals and Aytu BioScience) while WSB from Reddit focused more on sectors heavily hit by COVID-19, such as airlines and theme parks (Boeing, American Airlines, and Disney). In June 2020, WSB favored casinos and online betting stocks such as MGM Resorts and DraftKings. Compared to Discord, meme stock tickers such as GME, BB, AMC, NOK dominated the first month of 2021 in WSB, and this shows that stocks that receive the most mentions can differ significantly between platforms, even for the same month. It is also noticeable that meme stocks start to appear several months earlier in Reddit WSB. While Figure 1.1 shows that overall chatter on both platforms is similar in magnitude, GameStop first appeared as one of the most discussed stocks on Reddit WSB in October 2020, while it first appeared on Discord in January 2021. This suggests that WSB may be a better platform for finding a leading indicator for potential meme stocks among the two platforms.

1.4.3 Portfolio Construction

Since the Discord data in my paper provides information on which stocks investors are discussing in live group chat rooms for each day, I test whether an investment strategy of buying these stocks is profitable. From this empirical test, not only are we able to see if social media-savvy investors are discussing the right stocks and discussing good investment opportunities, but we are also able to observe whether live group interaction between investors helps make better choices. Since investors discuss through live group chat rooms, the data collected from the retail chatter is similar to aggregating discussions from offline investment clubs. Results can be compared to investors that look for investment opportunities by simply scrolling and reading posted messages from a different platform.

Popular stocks are collected from market close on day $t-1$ to market open on day t, the portfolio is rebalanced at the start of each trading day t, and OTO portfolio returns are computed from market open on day t to day $t+1$. For comparison, I also use monthly returns and factors to adjust the rebalancing interval and rebalance at the beginning of each month. Following this procedure, I construct the Discord portfolios that are investible in time.

1.4.4 Graphical Performance by Weighting Scheme

Figure 1.2 depicts how investing in a VW and EW portfolio would have evolved. Both portfolios are compared with the S&P 500. The technology sector got hit the most during the September selloff: during the first four trading days of September, Tesla dropped 30%, DocuSign dropped 24%, Apple dropped 16%, and CrowdStrike dropped 13%. Since the daily-rebalanced portfolios are constructed with only 50 stocks each day, a sharp drop is visible in that period since it contained these technology stocks during that period.

Results show that the S&P 500 underperformed compared to other portfolios. Regarding the weighting scheme, the EW portfolio outperformed for Discord in Panel A, suggesting that small and mid-cap popular stocks outperformed. In contrast, the VW portfolio performed better for Reddit WSB, suggesting that large-cap popular stocks performed well. Monthly rebalancing shows similar results. Overall, investing in the Discord EW portfolio would have been the most profitable among these portfolios.

1.4.5 Performance with respect to Benchmark Models

Table 1.8 analyzes the daily rebalancing portfolio return performance with respect to various benchmark models for both weighting schemes. The 0-F benchmark is the mean (net of the risk-free rate), the 1-F benchmark is the Capital Asset Pricing Model (CAPM), the 3-F benchmark is the Fama and French (1993) three-factor model, the 5-F benchmark is the Fama and French (2015) five-factor model, and the 6-F benchmark is the Fama and French (2015) five-factor model plus momentum.

Comparing alphas by weighting scheme, Panel A shows that the EW scheme outperformed for Discord, and the VW scheme outperformed for Reddit WSB in Panel B, consistent with the results in Figure 1.2. When comparing the magnitude of the alphas between portfolios, the Discord EW portfolio shows better abnormal performance on all models. In particular, the average daily performance of the Discord EW portfolio is much larger and statistically significant (37 bp per day) on the 0-F benchmark model than the Reddit WSB EW portfolio (11 bp per day) and the Reddit WSB VW portfolio (14 bp per day). Results further show that the Discord EW portfolio did not underperform, has little connection to the market, and has a tilt towards small-cap and growth stocks. This suggests that investors in Discord are willing to take the time to find and research small companies worth buying that have comparably less focus and attention to those of large-cap companies. Another explanation for this is that once stocks reach a high enough valuation, the influence popularity has on its future price is reduced since it has a larger market capitalization and therefore is harder to move the price.

In Table 1.9, I change the rebalancing interval by using monthly returns and factors and rebalance at the start of each month. The Discord EW portfolio offers an alpha of 6.1% per month using the 0-F model and 6.2% per month using the 1-F model with *t*-statistics of 2.45 and 1.99, respectively. Similar to Table 1.8, which uses daily rebalancing, the Discord EW portfolio shows better abnormal performance against nearly all models.

My paper also tests whether popular stocks are correlated with the performance of Bitcoin, as Bitcoin investing is often associated with fear of missing out, which is linked to undeserved popularity. In untabulated analysis, I find that Reddit WSB-popular stocks have nearly three times higher loading on Bitcoin returns than Discord-popular stocks. This is consistent with the Reddit WSB-popular stocks containing many meme stocks. If a social media platform has more discussions about meme stocks, following the crowd may lead to significant investment losses when the timing of jumping on and off the bandwagon is not timed well.

In sum, results show that Reddit WSB has much more and earlier mentions of meme stocks, the portfolio is less profitable, and is more correlated with Bitcoin returns. This suggests that Reddit WSB is more suitable for hedge funds that need to pay more attention to meme stocks gaining popularity to protect their portfolio. For focused quality discussion and self-growth, it may be beneficial for retail investors to monitor the Discord platform. The results from the portfolio analyses provide evidence that social media group investing can provide investment value and outperform the market if used the correct way.

1.4.6 Sentiment Analysis

Previous studies find that the vast majority of messages on online message boards represent buy signals (Dewally, 2003; Antweiler and Frank, 2004), implying that an increase in mentions should be associated with an increase in bullishness. Since popular stocks are the most discussed stocks, the portfolios are long-only in Sections 1.4.3 to 1.4.5. To see if tracking sentiment during portfolio daily rebalancing still has any impact, I consider the direction of the signals by opening a long position if the average sentiment of all discussions containing the given ticker is positive and short position otherwise. The sentiment is calculated using VADER, a lexicon and rule-based sentiment analysis tool specifically attuned to sentiments expressed in social media.[21] I set the classification threshold at −0.1 and +0.1 for all normalized sentiment scores between −1 and 1. Therefore, I classify positive sentiment if the score is greater than or equal to 0.1 and negative sentiment if the score is less than or equal to −0.1. I find a large imbalance between the two sentiments, where about 80% are classified as positive sentiment. This is consistent with previous studies that find that online stock advice is overwhelmingly

[21] For example, VADER works better than other sentiment analysis tools when sentences contain many slang words, emoticons, utf-8 encoded emojis, initialisms, and acronyms. See Hutto and Gilbert (2014) for details.

positive, with a ratio of buy advice to sell advice greater than 7:1 in Dewally (2003) and 5:1 in Antweiler and Frank (2004).

Figure 1.3 shows that, similar to Figure 1.2 Panel A, both EW and VW Long/Short portfolios outperformed the S&P 500, which is not surprising since most of the sentiment is classified as positive. It is still notable that the overall performance of both portfolios significantly improved. Nonetheless, it seems that the tone of the stock discussions is less of an issue than identifying "which" stocks investors are discussing the most since they are mostly positive in sentiment.

1.4.7 Possible Explanation for the Difference in Platforms

Several reasons may explain the different results between platforms. First, social media-savvy investors that search for investing servers in Discord and joining by an invitation link may have different incentives than investors that create accounts in Reddit and write in WSB. It is more likely that the investors joining Discord are searching for real investment advice and are planning to find a good group of people to share information. Live interactions among users may also create an incentive to recommend better plays, especially if users plan to stay in their specific investing group. In contrast, it may be harder to identify the intentions of posts and comments from random individuals on an asynchronous platform. As a result, when investors gather in chat rooms for the same goal of building wealth, the social media group investing crowd may often demonstrate wisdom.

Second, posts in the past do not matter much in the analysis since the main results are rebalanced daily. In other words, the only information that is relevant is the information available at the moment. Therefore, what is coming through as information right now is more likely better than what is getting posted as information. Monitoring by moderators may also help make the discussions more organized with minimal off-topic chats.

Lastly, through live discussions, information gets corrected more quickly, which results in a faster Bayesian updating among market participants. For example, if user X has incorrect information about stock A and user Y has the correct information about stock A, it is more likely that user X will update his belief with correct information faster during live discussions than from posts and comments on an asynchronous platform.

1.5 Conclusion

This paper investigates online group chatter related to investing. Specifically, I study whether social media group investing through live chat provides better returns by comparing a synchronous platform and an asynchronous platform. I also see whether social media group investing is related to future abnormal volume, stock volatility, and returns. As part of this analysis, I see whether the empirical results support the informed trading explanation or the price pressure explanation.

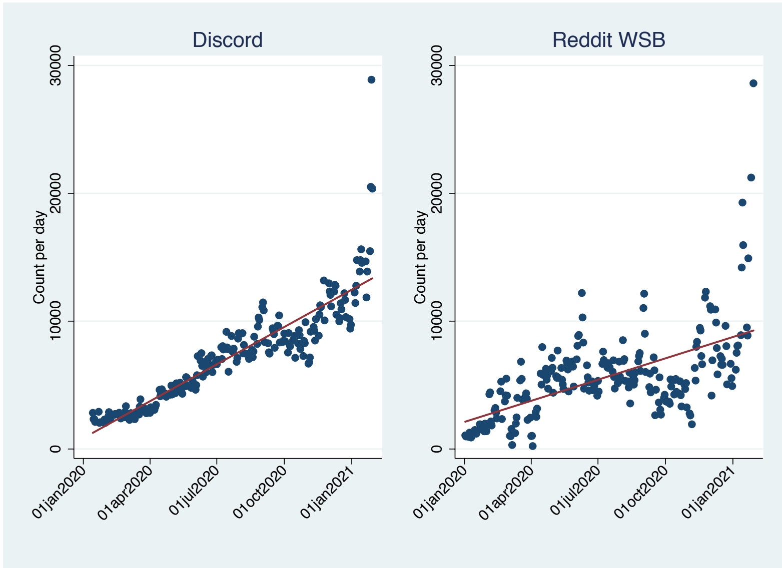
My paper is the first to examine Discord, a social media platform with group chatting capability and other distinctive features compared to other platforms. Using this unique data set, I create a direct measure for "popularity" that can also be used in other studies. The market began to truly acknowledge the impact of social media, which is why this topic is of interest and important for both retail and institutional investors. While most studies in the related literature use a broader measure of retail investors, this paper particularly examines social media-savvy retail investors. It focuses on where the smart chatter occurs, what type of stocks get more discussion, and whether the discussions are informed during a macroeconomic condition of market uncertainty.

Examining the impact of social media group investing, I find a positive relation between popularity and future abnormal trading volume, stock volatility, and returns. Subsample analysis shows that the impact is stronger for smaller firms. In addition, I find that higher returns are driven by continuity of popularity than by the day it becomes popular. I investigate channels for why more stock discussions may be associated with an increase in stock returns. I find that the buy-and-hold returns show a steady, continued upward drift over long horizons, supporting the informed trading explanation related to deserved popularity. I also find that insiders are less likely to sell their shares after a stock becomes popular, inconsistent with managerial opportunism, and lending further support for the informed trading explanation.

My paper presents evidence that social media-savvy investors discuss differently between synchronous and asynchronous platforms. Social media group investing through Discord provides investment value and can outperform the market. The Discord portfolio did not underperform with respect to the Fama-French benchmark models. The popular stocks tilt towards small growth stocks, suggesting that Discord's social media-savvy investors are willing to take the time to research and identify small companies worth buying that have less focus and attention than large-cap companies. In contrast, worse portfolio performance, and much more and earlier meme stock discussions suggest that Reddit WSB is better suited for institutional investors looking to protect their investments. Therefore, investors interested in generating higher alpha by identifying opportunities from social media should keep close track of Discord rather than Reddit. I also find that investor sentiment is less of an issue than identifying “which” stocks investors mostly discuss. The live group chat functionality with moderators, faster information with Bayesian updating in beliefs, and the users’ incentives may contribute to explaining Discord’s smarter chatter with better stock selections.

The trend of social media group investing continues to grow. The key findings in my paper suggest that it may be helpful for an investor to surround oneself with social media-savvy investors by jumping into these platforms because quite often, great ideas come from them. Moreover, adding a “Historical Mentions by Platform” feature may be valuable for online brokers to attract new investors and help them be more informed in their decision-making process. Lastly, the implications of popularity in my paper may hopefully serve as a basis for creating a popularity factor that may potentially become an essential factor in the future.

Panel A: Difference between Platforms



Panel B: Difference within Day

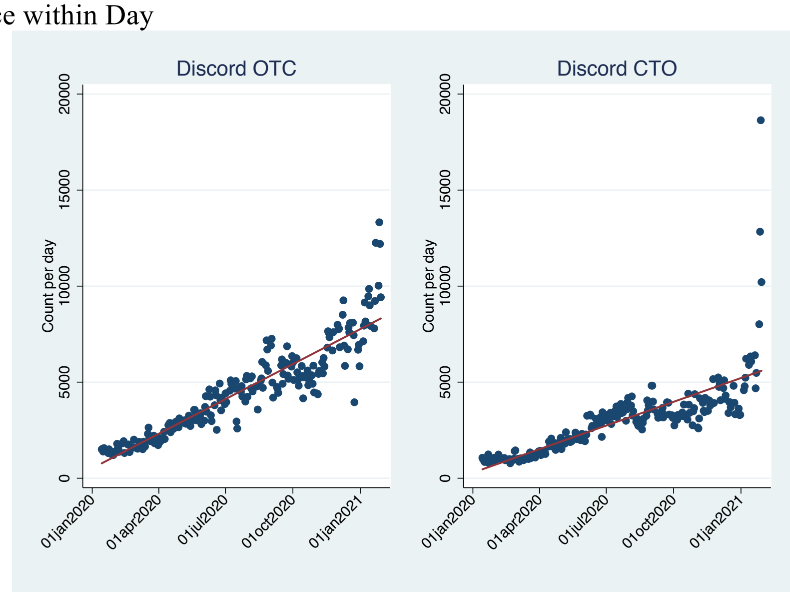
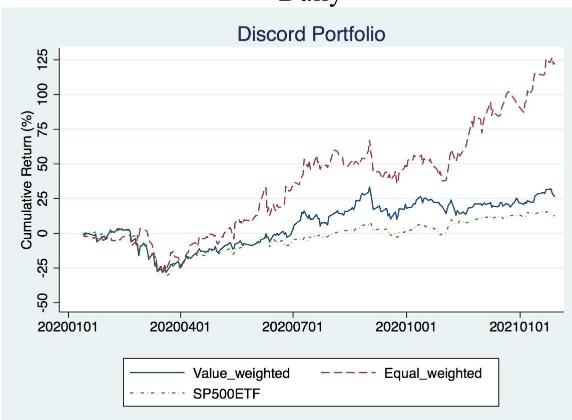


Figure 1.1: Top 50 Stock's Mention Count Over Time

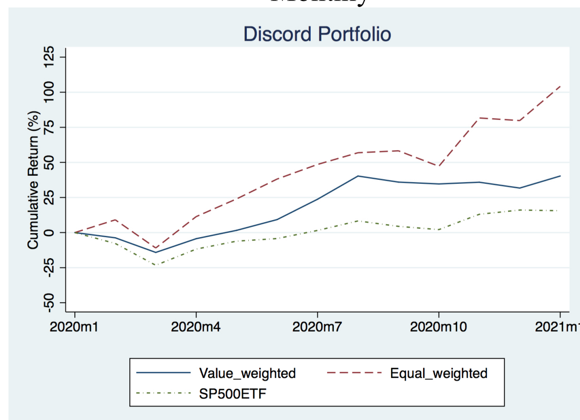
The figure shows the total count of the top 50 stocks most-mentioned each day from investing chat rooms in Discord and the Daily Discussion thread of WSB from Reddit. Panel A shows the results for the 24-hour collection period from Discord and Reddit WSB. Panel B divides the 24-hour period from Discord into two periods: market open on day t to market close on day t (OTC) and market close on day $t-1$ to market open on day t (CTO).

Panel A: Cumulative Returns of the Discord Portfolio

Daily

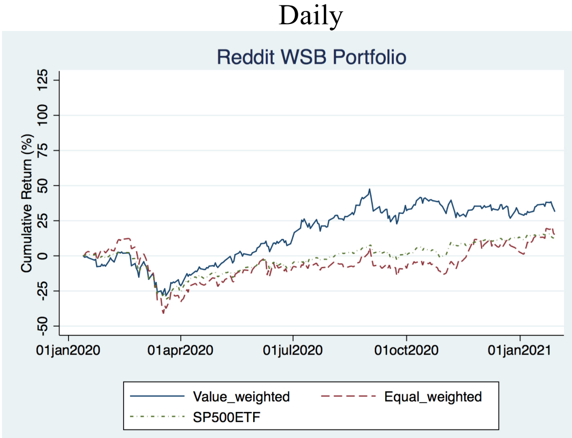


Monthly



Panel B: Cumulative Returns of the Reddit WSB Portfolio

Daily



Monthly

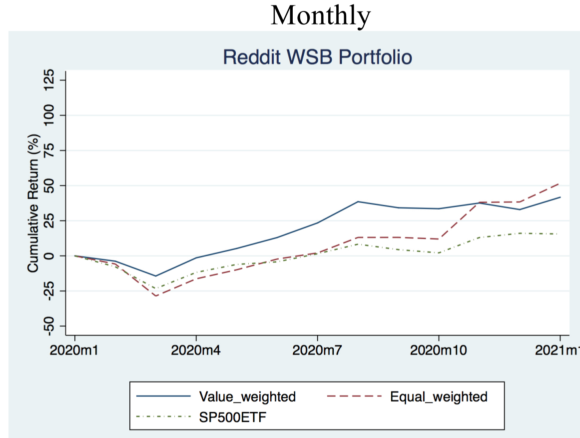


Figure 1.2: Portfolio Cumulative Returns by Weighting Scheme

The figure depicts how investing in a value-weighted and equal-weighted Popularity portfolio would have evolved from January 2020 to January 2021. Both portfolios are compared with the S&P 500 using daily and monthly rebalancing. Panel A shows the results for the Discord portfolio and Panel B shows the results for the Reddit WSB portfolio.

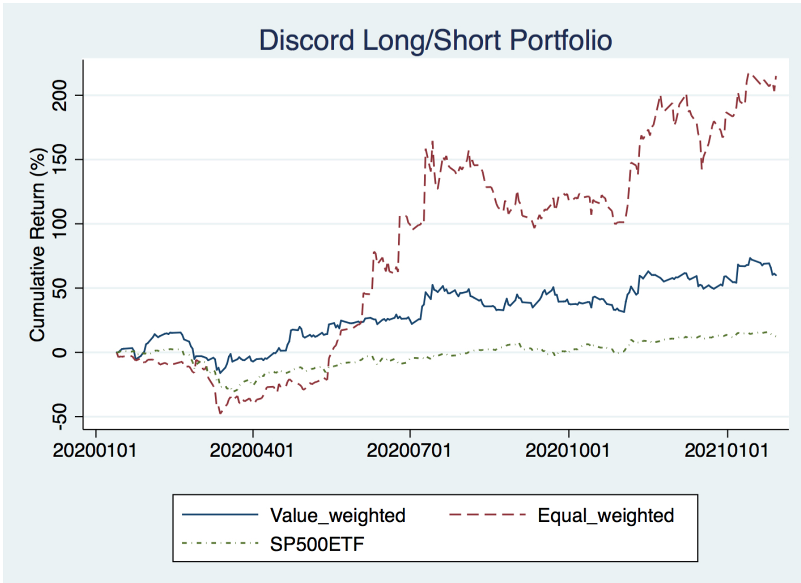


Figure 1.3: Portfolio Cumulative Returns using Sentiment Analysis

The figure depicts how investing in a value-weighted and equal-weighted Popularity Long/Short portfolio would have evolved from January 2020 to January 2021. The classification threshold is set at -0.1 and $+0.1$ for all normalized sentiment scores between -1 and 1 and is classified as positive sentiment if the score is greater than or equal to 0.1 and negative sentiment if the score is less than or equal to -0.1. The Discord portfolios are compared with the S&P 500. The portfolios are rebalanced at the start of each trading day t.

Table 1.1: Monotonic Relationships

This table reports the number of unique tickers in each rank tier, the average percent of days the stocks in each tier are included in the most-mentioned list during the sample period, and the average market capitalization in billion dollars. For each day from January 2020 to January 2021, popular stocks are sorted based on their mention rankings into five deciles (the highest tier of 1–10 to the lowest tier of 41–50) and aggregated across the sample period. Panel A shows the results for the 24-hour collection period from Discord and Reddit WSB. Panel B divides the 24-hour period from Discord into two periods: market open on day t to market close on day t (OTC) and market close on day $t-1$ to market open on day t (CTO).

Panel A: Difference between Platforms

Discord				Reddit WSB			
Tier	Unique Tickers	Avg Inclusion	Mkt Cap	Tier	Unique Tickers	Avg Inclusion	Mkt Cap
1-10	128	71.40%	529 B	1-10	138	65.07%	560 B
11-20	202	55.00%	355 B	11-20	219	47.56%	405 B
21-30	237	46.82%	351 B	21-30	251	38.12%	276 B
31-40	263	36.70%	304 B	31-40	283	32.12%	202 B
41-50	279	29.77%	191 B	41-50	305	27.38%	148 B

Panel B: Difference within Day

Discord OTC				Discord CTO			
Tier	Unique Tickers	Avg Inclusion	Mkt Cap	Tier	Unique Tickers	Avg Inclusion	Mkt Cap
1-10	130	74.19%	524 B	1-10	167	61.53%	472 B
11-20	199	58.52%	381 B	11-20	220	50.99%	398 B
21-30	228	46.46%	349 B	21-30	270	45.76%	341 B
31-40	272	35.30%	241 B	31-40	287	36.27%	287 B
41-50	279	29.05%	201 B	41-50	299	28.76%	247 B

Table 1.2: Popularity, Trading Volume, and Stock Volatility

This table presents results from daily Fama and MacBeth (1973) regressions of abnormal volume (AbVolume) and stock volatility (|AbRet|) at day *t* on popularity and other control variables from January 2020 to January 2021. Popular(t−1) is an indicator variable for the top 50 most-mentioned stocks on Discord from market close on day *t*−1 to market open on day *t* (CTO). Volume(t−1) is log of the trading volume on day *t*−1, Size(t−1) is log of market capitalization on day *t*−1, Ret[−5, −1] is cumulative returns from day *t*−5 through day *t*−1, Illiquidity[−5, −1] is the average of Amihud's (2002) illiquidity measure over days *t*−5 to *t*−1. Book-to-Market is the log of the ratio of book equity from the most recent fiscal year to the market capitalization at the end of the year, Profitability is operating income before depreciation as a fraction of average total assets based on most recent two periods, Volatility is the log of the standard deviation of daily returns from the previous month, and Momentum is the return over the past 12 months without the most recent month's return. Columns Small and Big in the table separately examine subsamples of below-median and above-median firm size for each day *t*. Newey and West (1987) adjusted *t*-statistics using 5 lags are reported in parenthesis.

Dependent Variable	AbVolume	AbVolume	AbVolume	|AbRet|	|AbRet|	|AbRet|
Firms Included	All	Small	Big	All	Small	Big
Popular(t−1)	0.231	0.485	0.196	0.013	0.035	0.007
	(15.87)	(9.40)	(15.65)	(6.30)	(4.38)	(8.44)
Volume(t−1)	0.027	0.035	0.019	0.003	0.004	0.002
	(13.61)	(14.32)	(10.78)	(13.79)	(13.79)	(13.97)
Size(t−1)	0.053	0.063	0.046	−0.003	−0.003	−0.003
	(16.99)	(11.81)	(16.10)	(−11.29)	(−8.58)	(−12.01)
Ret[−5, −1]	0.111	0.115	−0.002	0.018	0.018	0.013
	(3.26)	(3.28)	(−0.04)	(5.02)	(4.82)	(3.38)
Illiquidity[−5, −1]	0.070	0.078	0.066	0.001	0.002	−0.000
	(25.50)	(23.05)	(23.44)	(7.63)	(10.78)	(−0.51)
Volatility	−0.050	−0.052	−0.051	0.007	0.008	0.007
	(−9.47)	(−7.01)	(−11.02)	(17.58)	(18.09)	(12.44)
Momentum	0.004	0.003	0.005	0.001	0.001	0.001
	(2.51)	(2.09)	(2.44)	(5.86)	(5.41)	(6.41)
Book−to−Market	−0.006	−0.007	−0.004	−0.0003	−0.000	−0.000
	(−2.95)	(−2.47)	(−2.61)	(−2.29)	(−1.02)	(−0.93)
Profitability	0.016	0.023	0.015	−0.005	−0.004	−0.006
	(2.05)	(2.95)	(1.35)	(−4.81)	(−3.75)	(−6.53)
N	445,484	226,458	219,026	445,383	226,376	219,007
R2	4.98%	5.84%	5.10%	13.61%	12.95%	13.49%

Table 1.3: Popularity and Future Stock Returns

This table presents results from daily Fama and MacBeth (1973) regressions of Ret[a,b] on popularity and other control variables from January 2020 to January 2021. Ret[a,b] is the OTO return in equation (1) compounded from day *t+a* to *t+b*. Popular(t−1) is an indicator variable for the top 50 most-mentioned stocks on Discord from market close on day *t*−1 to market open on day *t* (CTO). AbVolume(t−1) is the lag of log trading volume on day *t* minus its average log trading volume on days *t*−5 to *t*−1, Size(t−1) is log of market capitalization on day *t*−1, Ret[−5, −1] is cumulative returns from day *t*−5 through day *t*−1, Illiquidity[−5, −1] is the average of Amihud's (2002) illiquidity measure over days *t*−5 to *t*−1. Book-to-Market is the log of the ratio of book equity from the most recent fiscal year to the market capitalization at the end of the year, Profitability is operating income before depreciation as a fraction of average total assets based on most recent two periods, Volatility is the log of the standard deviation of daily returns from the previous month, and Momentum is the return over the past 12 months without the most recent month's return. Newey and West (1987) adjusted *t*-statistics using lags up to two times the holding period are reported in parenthesis.

Dependent Variable	Ret[0,0]	Ret[1,1]	Ret[1,5]	Ret[1,30]	Ret[1,45]	Ret[1,60]
Firms Included	All	All	All	All	All	All
Popular(t−1)	0.003	0.003	0.006	0.040	0.053	0.066
	(1.63)	(2.06)	(1.68)	(3.12)	(3.94)	(2.72)
AbVolume(t−1)	0.000	−0.000	0.000	0.001	0.001	0.004
	(1.29)	(−0.01)	(0.52)	(0.60)	(0.68)	(2.55)
Size (t−1)	−0.001	−0.001	−0.004	−0.028	−0.043	−0.056
	(−2.76)	(−3.17)	(−3.68)	(−3.50)	(−4.14)	(−7.55)
Ret[−5, −1]	−0.009	−0.003	−0.012	−0.000	0.004	0.029
	(−1.94)	(−0.56)	(−0.63)	(−0.00)	(0.13)	(0.59)
Illiquidity[−5, −1]	−0.001	−0.001	−0.003	−0.018	−0.028	−0.036
	(−2.25)	(−2.40)	(−3.32)	(−3.40)	(−4.19)	(−7.43)
Volatility	0.002	0.002	0.010	0.061	0.092	0.117
	(2.46)	(2.79)	(2.97)	(3.61)	(5.11)	(9.75)
Momentum	−0.000	−0.000	−0.000	−0.003	−0.007	−0.013
	(−0.45)	(−0.79)	(−0.44)	(−1.00)	(−2.35)	(−6.31)
Book−to−Market	−0.000	−0.000	−0.002	−0.007	−0.007	−0.006
	(−0.95)	(−1.01)	(−1.63)	(−0.88)	(−0.53)	(−0.33)
Profitability	−0.002	−0.002	−0.009	−0.040	−0.028	−0.023
	(−1.62)	(−1.75)	(−2.07)	(−1.31)	(−0.77)	(−0.51)
N	445,320	445,230	444,872	442,759	441,556	440,330
R2	11.23%	11.08%	11.54%	11.01%	12.10%	12.86%

Table 1.4: Entry and Continuity of Popularity and Future Stock Returns

This table presents results from daily Fama and MacBeth (1973) regressions of Ret[a,b] on Entry(t−1), Continuity(t−1), and other control variables from January 2020 to January 2021. Ret[a,b] is the OTO return in equation (1) compounded from day $t+a$ to $t+b$. Entry(t−1) is an indicator variable that equals one if the stock was not in the Discord most-mentioned list on day $t-2$ and entered the Discord most-mentioned list on day $t-1$. Continuity(t−1) is an indicator variable that equals one if the stock was in the Discord most-mentioned list on day $t-2$ and also on day $t-1$. AbVolume(t−1) is the lag of log trading volume on day t minus its average log trading volume on days $t-5$ to $t-1$, Size(t−1) is log of market capitalization on day $t-1$, Ret[−5, −1] is cumulative returns from day $t-5$ through day $t-1$, Illiquidity[−5, −1] is the average of Amihud's (2002) illiquidity measure over days $t-5$ to $t-1$. Book-to-Market is the log of the ratio of book equity from the most recent fiscal year to the market capitalization at the end of the year, Profitability is operating income before depreciation as a fraction of average total assets based on most recent two periods, Volatility is the log of the standard deviation of daily returns from the previous month, and Momentum is the return over the past 12 months without the most recent month's return. Newey and West (1987) adjusted t-statistics using lags up to two times the holding period are reported in parenthesis.

Dependent Variable	Ret[0,0]	Ret[1,1]	Ret[1,5]	Ret[1,30]	Ret[1,45]	Ret[1,60]
Firms Included	All	All	All	All	All	All
Entry(t−1)	0.001	−0.000	0.005	0.027	0.040	0.059
	(0.35)	(−0.14)	(1.04)	(2.15)	(2.68)	(2.12)
Continuity(t−1)	0.004	0.003	0.006	0.041	0.054	0.066
	(1.68)	(2.09)	(1.53)	(4.07)	(4.41)	(3.04)
AbVolume(t−1)	0.000	0.000	0.000	0.001	0.001	0.004
	(1.29)	(0.07)	(0.52)	(0.67)	(0.74)	(2.59)
Size (t−1)	−0.001	−0.001	−0.004	−0.028	−0.043	−0.056
	(−2.73)	(−3.18)	(−3.67)	(−3.49)	(−4.13)	(−7.54)
Ret[−5, −1]	−0.009	−0.003	−0.012	−0.001	0.004	0.028
	(−1.94)	(−0.55)	(−0.61)	(−0.01)	(0.13)	(0.57)
Illiquidity[−5, −1]	−0.001	−0.001	−0.003	−0.018	−0.028	−0.036
	(−2.21)	(−2.40)	(−3.31)	(−3.40)	(−4.18)	(−7.41)
Volatility	0.000	0.002	0.010	0.061	0.092	0.117
	(1.29)	(2.79)	(2.96)	(3.61)	(5.12)	(9.79)
Momentum	−0.000	−0.000	−0.000	−0.003	−0.008	−0.013
	(−0.39)	(−0.80)	(−0.44)	(−1.01)	(−2.38)	(−6.32)
Book−to−Market	−0.000	−0.000	−0.002	−0.007	−0.007	−0.006
	(−0.94)	(−1.01)	(−1.63)	(−0.88)	(−0.53)	(−0.33)
Profitability	−0.002	−0.002	−0.009	−0.040	−0.028	−0.023
	(−1.62)	(−1.73)	(−2.06)	(−1.31)	(−0.77)	(−0.51)
N	445,320	445,230	444,872	442,759	441,556	440,330
R2	11.45%	11.19%	11.66%	11.10%	12.18%	12.93%

Table 1.5: Insiders' Incentives and Opportunities

This table presents results from pooled panel regressions of InsiderSell and InsiderNSR on $\text{Day}[a, b]$ and other control variables from January 2020 to January 2021. InsiderSell is an indicator variable equal to one if insiders at the firm are net sellers on that day and zero otherwise. InsiderNSR is the ratio of net sales to total insider transactions. $\text{Day}[a, b]$ is an indicator variable that equals one if the day falls within the window and zero otherwise. $\text{Day}[0]$ is an indicator variable that equals one if the stock was not in the Discord most-mentioned list the previous day and entered the list on day *t*=0. The same control variables from Tables 6-8 are used. Columns differ in terms of whether firm-month fixed effects with clustered standard errors are included and whether overlapping windows are allowed.

Dependent Variable	InsiderSell	InsiderSell	InsiderSell	InsiderNSR	InsiderNSR	InsiderNSR
Day[−10,−1]	−0.021	−0.045	−0.012	−0.041	−0.089	−0.026
	(−1.61)	(−2.65)	(−0.88)	(−1.57)	(−2.62)	(−1.00)
Day[0]	−0.091	−0.103	−0.005	−0.185	−0.207	−0.014
	(−3.79)	(−3.29)	(−0.29)	(−3.86)	(−3.34)	(−0.44)
Day[+1,+10]	−0.083	−0.104	−0.042	−0.168	−0.210	−0.085
	(−6.47)	(−6.64)	(−2.47)	(−6.59)	(−6.74)	(−2.50)
N	23,416	22,897	22,742	23,416	22,897	22,742
R2	20.58%	21.08%	63.45%	20.79%	21.30%	63.84%
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Yes	No	No	Yes
Window Overlap	Yes	No	No	Yes	No	No

Table 1.6: Overall Most-mentioned List

This table shows the overall top 15 most-mentioned stocks from January 2020 to January 2021 for Discord and Reddit WSB. The Global Industry Classification Standard (GICS) is shown in the middle column. The percentage column represents the total number of times a stock ticker is mentioned divided by the total number of times the top 50 stocks are mentioned during the sample period. Not shown, the majority of tickers that represent less than 1% differ greatly between platforms.

Discord			Reddit WSB		
Ticker	GICS Classification	Percent	Ticker	GICS Classification	Percent
TSLA	Automobile Manufacturers	6.98%	TSLA	Automobile Manufacturers	10.10%
AAPL	Technology Hardware	4.62%	GME	Specialty Retail	8.58%
BA	Aerospace & Defense	3.07%	PLTR	Application Software	4.75%
AMD	Semiconductors	2.82%	AAPL	Technology Hardware	4.39%
NIO	Automobile Manufacturers	2.67%	BB	Software	3.69%
AMZN	Internet Retail	2.19%	AMD	Semiconductors	3.60%
FB	Interactive Media & Services	2.10%	NIO	Automobile Manufacturers	3.48%
ZM	Application Software	2.02%	MSFT	Systems Software	3.11%
MSFT	Systems Software	1.84%	AMC	Movies & Entertainment	2.94%
GME	Specialty Retail	1.70%	BA	Aerospace & Defense	2.56%
BABA	Internet Retail	1.70%	AMZN	Internet Retail	2.22%
NFLX	Movies & Entertainment	1.60%	NOK	Communications Equipment	1.99%
ROKU	Movies & Entertainment	1.46%	SPCE	Aerospace & Defense	1.92%
NVDA	Semiconductors	1.36%	BABA	Internet Retail	1.87%
SPCE	Aerospace & Defense	1.31%	NKLA	Machinery & Heavy Trucks	1.66%

Table 1.7: Top 5 Most Mentions by Month

This table shows the top five most-mentioned stocks for each month from January 2020 to January 2021. Panel A shows the results for Discord and Panel B shows the results for Reddit WSB.

Panel A: Most-mentioned Stocks on Discord by Month

Month Year	Rank 1	Rank 2	Rank 3	Rank 4	Rank5
2020m1	TSLA	BYND	AAPL	BA	NFLX
2020m2	TSLA	SPCE	AAPL	MSFT	AMD
2020m3	BA	TSLA	AAPL	INO	AYTU
2020m4	TSLA	BA	AMD	NFLX	AAPL
2020m5	MARK	TSLA	BA	AAPL	AMD
2020m6	BA	TSLA	IDEX	GNUS	AAPL
2020m7	TSLA	BA	AAPL	AMD	AMZN
2020m8	TSLA	AAPL	BA	MSFT	FB
2020m9	TSLA	AAPL	ZM	NKLA	AMD
2020m10	AAPL	TSLA	AMD	NIO	ZM
2020m11	NIO	TSLA	ZM	BABA	AAPL
2020m12	TSLA	AAPL	PLTR	NIO	AMD
2021m1	GME	AMC	TSLA	AAPL	NIO

Panel B: Most-mentioned Stocks on Reddit WSB by Month

Month Year	Rank 1	Rank 2	Rank 3	Rank 4	Rank5
2020m1	TSLA	SPCE	AMD	BA	MSFT
2020m2	SPCE	TSLA	MSFT	AMD	DIS
2020m3	BA	TSLA	DIS	AAL	AMD
2020m4	TSLA	AMD	BA	MSFT	SNAP
2020m5	TSLA	DIS	DKNG	FB	AMD
2020m6	BA	MGM	NKLA	DKNG	TSLA
2020m7	TSLA	AMZN	MSFT	AMD	WMT
2020m8	TSLA	AAPL	RKT	MSFT	PRPL
2020m9	TSLA	AAPL	NKLA	RKT	PTON
2020m10	AAPL	AMD	TSLA	GME	NIO
2020m11	NIO	PLTR	TSLA	BABA	RKT
2020m12	PLTR	GME	TSLA	NIO	BABA
2021m1	GME	BB	AMC	NOK	PLTR

Table 1.8: Portfolio Performance to Benchmark Models Daily Rebalancing

This table analyzes the daily rebalancing return performance of the Popularity portfolios with respect to various benchmark models for both weighting schemes from January 2020 to January 2021. The 0-F benchmark is the mean (net of the risk-free rate), 1-F benchmark is the Capital Asset Pricing Model (CAPM), 3-F benchmark is the Fama and French (1993) three-factor model, 5-F benchmark is the Fama and French (2015) five-factor model, and 6-F benchmark is the Fama and French (2015) five-factor model plus momentum. Panel A shows the results for the Discord portfolio and Panel B shows the results for the Reddit WSB portfolio. The portfolios are rebalanced at the start of each trading day *t*. The alpha is expressed as a percentage. Newey and West (1987) adjusted *t*-statistics using 1 lag is reported in parenthesis.

Panel A: Performance of the Discord portfolio

	Equal-weighted					Value-weighted				
	O-F	1-F	3-F	5-F	6-F	O-F	1-F	3-F	5-F	6-F
ALPHA	0.370	0.352	0.214	0.236	0.235	0.121	0.102	0.000	0.001	−0.01
	(2.03)	(2.05)	(1.30)	(1.44)	(1.42)	(0.88)	(0.79)	(0.00)	(0.01)	(−0.09)
XMKT		0.168	0.225	0.203	0.204		0.183	0.297	0.281	0.291
		(0.93)	(1.26)	(1.08)	(1.13)		(1.24)	(2.02)	(1.87)	(2.04)
SMB			1.117	0.957	0.956			0.325	0.354	0.342
			(4.22)	(3.31)	(3.36)			(1.65)	(1.75)	(1.73)
HML			−0.461	−0.51	−0.52			−0.599	−0.722	−0.802
			(−2.80)	(−2.09)	(−1.74)			(−6.27)	(−4.96)	(−3.89)
RMW				−0.567	−0.572				0.400	0.361
				(−1.39)	(−1.36)				(1.53)	(1.38)
CMA				−0.337	−0.334				−0.151	−0.133
				(−0.78)	(−0.78)				(−0.57)	(−0.50)
UMD					−0.01					−0.077
					(−0.05)					(−0.55)

Panel B: Performance of the Reddit WSB portfolio

	Equal-weighted					Value-weighted				
	O-F	1-F	3-F	5-F	6-F	O-F	1-F	3-F	5-F	6-F
ALPHA	0.107	0.089	−0.054	−0.040	−0.031	0.137	0.115	0.018	0.020	0.016
	(0.60)	(0.54)	(−0.36)	(−0.27)	(−0.20)	(0.96)	(0.88)	(0.14)	(0.16)	(0.13)
XMKT		0.196	0.244	0.212	0.204		0.225	0.332	0.313	0.316
		(0.97)	(1.22)	(1.01)	(0.98)		(1.51)	(2.22)	(2.02)	(2.13)
SMB			1.280	1.214	1.224			0.337	0.362	0.358
			(4.57)	(4.22)	(4.34)			(1.62)	(1.68)	(1.69)
HML			−0.437	−0.607	−0.541			−0.566	−0.672	−0.697
			(−2.97)	(−2.76)	(−1.92)			(−5.65)	(−4.33)	(−3.29)
RMW				−0.222	−0.189				0.362	0.350
				(−0.60)	(−0.50)				(1.36)	(1.30)
CMA				−0.382	−0.397				−0.226	−0.22
				(−0.92)	(−0.97)				(−0.80)	(−0.79)
UMD					0.064					−0.024
					(0.39)					(−0.17)

Table 1.9: Portfolio Performance to Benchmark Models Monthly Rebalancing

This table analyzes the monthly rebalancing return performance of the Popularity portfolios with respect to various benchmark models for both weighting schemes from January 2020 to January 2021. The 0-F benchmark is the mean (net of the risk-free rate), 1-F benchmark is the Capital Asset Pricing Model (CAPM), 3-F benchmark is the Fama and French (1993) three-factor model, 5-F benchmark is the Fama and French (2015) five-factor model, and 6-F benchmark is the Fama and French (2015) five-factor model plus momentum. Panel A shows the results for the Discord portfolio and Panel B shows the results for the Reddit WSB portfolio. The portfolios are rebalanced at the start of each month. The alpha is expressed as a percentage. Newey and West (1987) adjusted *t*-statistics using 1 lag is reported in parenthesis.

Panel A: Performance of the Discord portfolio

	Equal-weighted					Value-weighted				
	O-F	1-F	3-F	5-F	6-F	O-F	1-F	3-F	5-F	6-F
ALPHA	6.127	6.221	2.910	4.120	4.939	2.543	2.322	0.099	0.875	−0.097
	(2.45)	(1.99)	(0.36)	(0.42)	(0.39)	(1.06)	(0.95)	(0.02)	(0.20)	(−0.02)
XMKT		−0.05	0.190	0.067	−0.002		0.116	0.355	0.303	0.384
		(−0.12)	(0.28)	(0.06)	(−0.00)		(0.47)	(1.34)	(0.67)	(0.77)
SMB			0.291	0.313	0.051			−0.047	−0.047	0.264
			(0.34)	(0.32)	(0.03)			(−0.05)	(−0.04)	(0.19)
HML			−0.97	−1.095	−1.013			−0.745	−0.858	−0.955
			(−0.57)	(−0.53)	(−0.42)			(−0.98)	(−0.89)	(−0.93)
RMW				1.200	1.272				0.884	0.799
				(0.52)	(0.47)				(0.42)	(0.36)
CMA				1.608	1.807				1.596	1.359
				(0.68)	(0.77)				(1.17)	(0.95)
UMD					−0.255					0.303
					(−0.23)					(0.51)

Panel B: Performance of the Reddit WSB portfolio

	Equal-weighted					Value-weighted				
	O-F	1-F	3-F	5-F	6-F	O-F	1-F	3-F	5-F	6-F
ALPHA	3.902	3.631	3.869	4.971	5.258	2.736	2.712	0.269	1.096	0.325
	(1.26)	(0.99)	(0.46)	(0.53)	(0.42)	(1.25)	(1.12)	(0.05)	(0.24)	(0.06)
XMKT		0.142	0.154	−0.060	−0.084		0.012	0.250	0.188	0.252
		(0.33)	(0.23)	(−0.06)	(−0.07)		(0.05)	(0.86)	(0.39)	(0.45)
SMB			−0.110	0.226	0.135			0.024	0.046	0.293
			(−0.11)	(0.25)	(0.07)			(0.03)	(0.04)	(0.22)
HML			0.035	−0.261	−0.232			−0.789	−0.947	−1.024
			(0.02)	(−0.13)	(−0.10)			(−0.92)	(−0.91)	(−0.88)
RMW				2.019	2.044				0.991	0.923
				(0.91)	(0.82)				(0.49)	(0.42)
CMA				2.432	2.502				1.793	1.605
				(0.87)	(0.90)				(1.27)	(1.07)
UMD					−0.089					0.240
					(−0.08)					(0.40)

Chapter 2

The Empirical Performance of $\mathbf{SVIX}_{i,t}^{2}$ as a Proxy for Expected Stock Returns

Joint work with Zhongjin Lu

2.1 Introduction

"...our model makes predictions about the quantitative relationship between expected returns and risk-neutral variances, we hope also to find that the estimated coefficients on the predictor variables are close to specific numbers that come out of the theory."

– Martin and Wagner (2019, emphasis in original)

An ex-ante measure of expected stock returns is highly desirable but extremely challenging to compute. Martin and Wagner (2019) propose to use one-half of the risk-neutral excess stock variance as a real-time proxy for expected excess-of-market stock returns. Unlike the other existing measures of expected stock returns, the expected return prediction in Martin and Wagner (2019) (MW model hereafter) specifies the predictor and the predictive coefficient based on a priori reasoning, which greatly minimizes the data mining concern. The importance of this advantage cannot be understated.

The goal of this paper is to help researchers have a thorough understanding of the sensitivity of the key empirical results in Martin and Wagner (2019) to reasonable variations in the empirical design. First, we correct a look-ahead bias introduced by a seemingly innocuous data filter that biases the estimated predictive coefficients in the original paper upward to the model-implied value of 0.5. Second, while Martin and Wagner (2019) use OLS panel regressions to estimate the predictive regression coefficients, we analyze whether the weighted-least-squares (WLS hereafter) and the Fama and MacBeth (1973) (FM hereafter) regressions yield substantially different coefficient estimates. Third, we extend the sample period by six years to December 2020 to examine the out-of-sample (OOS) model performance in the post-publication period and see how it evolves over time. Fourth, we examine the MW model's performance in subsamples sorted on popular stock characteristics that determine popular equity investment styles. Finally, we investigate the reasons why the MW model performs well in some periods but not in others.

For ease of comparison, we benchmark all our results to the key results in Martin and Wagner (2019) from running a pooled OLS panel regression of excess-of-market returns of S&P 500 stocks on excess stock variance for various forecasting horizons.[22] They find a predictive regression coefficient ranging from 0.560 to 0.949 with firm fixed effects and from 0.301 to 0.553 without firm fixed effects. In all cases, they cannot reject the null hypothesis that the predictive regression coefficient is equal to 0.5 at the 5% level.

[22] Martin and Wagner (2019) find similar results using S&P 100 firms. Our results using S&P 100 firms are available upon request.

We find that after correcting for the look-ahead bias, the pooled OLS regression coefficients shrink by about 16% to 18% and 26% to 34% with and without firm fixed effects, respectively, across forecasting horizons. Furthermore, in both the original and our extended sample periods, if we use the WLS and FM regressions, the estimated regression coefficients are 64% to 185% smaller than the pooled OLS coefficients in all but the 1-month forecasting horizon using WLS. In all specifications, the WLS and FM regression coefficients are not statistically significantly different from zero for all forecasting horizons.

The finding that the WLS and FM regression coefficients are substantially lower than the pooled OLS regression coefficients is related to the fact that the panel regressions give more weight to periods with larger cross-sectional variation in the predictor than the WLS and FM regressions. When we re-estimate the panel regressions with the exclusion of 5% of the months with the largest cross-sectional standard deviation in the predictor, we find similar patterns that the resulting predictive regression coefficients are substantially lower and become statistically insignificant in all cases.

Consistent with these in-sample analyses, our OOS analysis (following the methodology of Welch and Goyal (2008)) shows that the superior performance of the MW model relative to the benchmark models comes almost entirely from the second half of the 2008-2009 financial crisis. Our subsample analysis shows that the predictive power of the risk-neutral variance for future stock returns is several times larger among value, unprofitable, small-cap, and past-loser stocks than among growth, profitable, large-cap, and past-winner stocks. The OOS plots for these different style subsamples show similar patterns that the superior out-of-sample performance of the MW model comes from the second half of the financial crisis. We explore several reasons for the time-varying performance of the MW model in Section 2.2.5.

2.2 Empirical Tests

2.2.1 Data

Our sample period is from January 4, 1996, through December 31, 2020. Following Martin and Wagner (2019), we gather the strike price and the option premium for S&P 500 stocks from the volatility surface files in OptionMetrics (OM). Forward prices are obtained from the standardized options data file in OM. Stock returns are provided by CRSP for the same period. Financial ratios are obtained from the Financial Ratios Suite by WRDS. Option-implied betas are obtained from Adrian Buss and Grigory Vilkov's website. We construct the individual stock risk-neutral variance $\text{SVIX}^2_{i,t,t+T}$ following Martin and Wagner (2019).[23] The $\text{SVIX}^2_{i,t,t+T}$ for the return horizon (which is also the option maturity) of 12 months is available on the Journal of Finance website and has a correlation of 0.9982 with our replicated $\text{SVIX}^2_{i,t,t+T}$. The average stock variance is then measured as the value-weighted sum of individual stock risk-neutral variances, $\overline{\text{SVIX}}^2_{t,t+T} = \sum_i w_{i,t} \text{SVIX}^2_{i,t,t+T}$. Our replicated excess stock variance $\text{SVIX}^2_{i,t,t+T} - \overline{\text{SVIX}}^2_{t,t+T}$ has a correlation of 0.9981 with their measure.

[23] At the publication of Martin and Wagner (2019), OptionMetrics provide only 13 OTM options in their "Standardized Options" file. In the latest release of OptionMetrics data, there are 17 OTM options available. We follow the original paper and use only 13 OTM options. We exclude the four most OTM options, which tend to have larger bid/ask spreads. See Internet Appendix A for more details.

Martin and Wagner (2019) require the stocks to remain as the S&P 500 index constituents for the entire forecasting horizon in their regression analyses. For example, with a forecasting horizon of 12 months, this data filter excludes stock-month observations 12 months before a stock is deleted from the S&P 500 index. Since knowing which stock will be deleted from the S&P 500 index in 12 months requires ex-post knowledge, this data filter introduces a look-ahead bias. We refer to this data filter as the MW data filter and conduct the analysis to quantify the resulting look-ahead bias.

2.2.2 Predictive Regressions

Our objective is to evaluate the sensitivity of the coefficient estimate in the following predictive regression to variations in specifications. At the end of each month t, we regress monthly overlapping future excess-of-market returns on excess risk-neutral stock variance,

$$\frac{E_t R_{i,t,t+T} - R_{m,t,t+T}}{R_{f,t,t+T}} = \alpha_i + \gamma\left(\mathrm{SVIX}^2_{i,t,t+T} - \overline{\mathrm{SVIX}}^2_{t,t+T}\right) + \epsilon_{i,t,t+T} \quad (1)$$

where $\mathrm{SVIX}^2_{i,t,t+T}$ is stock i's risk-neutral variance at time t for a horizon of T, $\overline{\mathrm{SVIX}}^2_{t,t+T}$ is the value-weighted average of $\mathrm{SVIX}^2_{i,t,t+T}$ across all stocks at time t, and $\mathrm{SVIX}^2_{i,t,t+T} - \overline{\mathrm{SVIX}}^2_{t,t+T}$ is the excess stock variance. For Tables 2.1 and 2.2, T is equal to 12, 6, 3, and 1 month in Panels A to D, respectively. We compute standard errors and p-values by employing the block bootstrap procedure in Martin and Wagner (2019) that is designed to take into account the time-series and cross-sectional dependencies in the data.[24]

[24] Specifically, Martin and Wagner (2019) use an overlapping block resampling scheme to estimate the covariance matrix of the estimated coefficients from 1,000 bootstrap samples. More details are provided in the Appendix

We start by focusing on T of 12 months in Panel A. Column (1) of Tables 2.1 and 2.2 report our replication of the pooled OLS panel regression results reported in Martin and Wagner (2019), without and with firm fixed effects, respectively. We use the same sample period and impose the same data filter of excluding stock-month observations 12 months before a stock is deleted from the S&P 500 index as in Martin and Wagner (2019). Column (1) of Panel A in Table 2.1 shows that when Eq. (1) is estimated without firm fixed effects, that is, constraining α_i to be the same across stocks, the predictive regression coefficient on our replicated excess stock variance is 0.54, which is almost identical to the corresponding estimate in Martin and Wagner (2019) (p.1907). In Table 2.2, when we include firm fixed effects, the predictive regression coefficient on our replicated excess stock variance is 0.87, which is very close to the corresponding coefficient of 0.92 in Martin and Wagner (2019) (p.1908). Like Martin and Wagner (2019), we find that the p-value for the hypothesis that $\gamma = 0.5$ cannot be rejected with and without fixed effects, whereas the hypothesis that $\gamma = 0$ can be rejected at the 1% level and 10% level with and without fixed effects, respectively. In Panels B to D for T equal to 6, 3, and 1 month, we find similar results that the hypothesis of $\gamma = 0.5$ cannot be rejected at the conventional levels. We refer to these OLS regression coefficients with and without firm fixed effects as the MW estimates and use them as the benchmark for comparison below.

In Column (2) of Tables 2.1 and 2.2, we examine the sensitivity of the predictive regression coefficient to the MW data filter that introduces the look-ahead bias. That is, Column

---

Section B1 of their paper. We acknowledge using the Martin and Wagner (2019) code posted on the Journal of Finance website for computing the bootstrap standard errors.

(2) keeps the same sample period as Martin and Wagner (2019) (like Column 1) but removes the MW data filter so that stocks are excluded from the regression only if they have been deleted from the S&P 500 index at month t. Consequently, the number of observations increases by about 6%. Column (2) of Panel A of Tables 2.1 and 2.2 shows that removing the MW data filter yields a predictive regression coefficient of 0.352 and 0.714 without and with firm fixed effects, respectively, 34% and 18% smaller than the corresponding MW estimates in Column (1). Column (2) of Panels B–D of Tables 2.1 to 2.2 repeat the above analyses for forecasting horizons of 6, 3, and 1 month and find a similar reduction in the regression coefficients: after removing the MW data filter, the pooled OLS regression coefficients shrink by about 16% to 18% across forecasting horizons with firm fixed effects and 26% to 34% without firm fixed effects. To avoid the effect of the look-ahead bias, we do not impose the MW data filter in the remaining columns.

In Column (3) of Tables 2.1 and 2.2, we examine the sensitivity of the predictive regression coefficient to the extension of the sample period to the end of 2020. Comparing Columns (2) and (3) of Tables 2.1 and 2.2, we find only small changes in the regression coefficients: the predictive regression coefficients slightly decrease for 6- and 12-month forecasting horizons but slightly increase for 1- and 3-month horizons.

In Columns (5) and (6) of Tables 2.1 and 2.2, we examine the sensitivity of the predictive regression coefficient to the regression methodology by using the same regression specifications in Columns (2) and (3) with the only change of using the FM and WLS regressions rather than the pooled OLS panel regressions used in Martin and Wagner (2019). We compare the FM regression coefficient with the OLS coefficient without firm fixed effects. For the WLS regression, we model the heteroscedasticity by regressing the absolute value of the residuals from the corresponding OLS regression on $\mathrm{SVIX}_{i,t,t+T}^2 - \overline{\mathrm{SVIX}}_{t,t+T}^2$. We then use the inverse of the fitted values as the weights in the WLS regression. We include firm fixed effects when running the WLS regression and therefore compare the WLS regression coefficient with the OLS coefficient with firm fixed effects.

Columns (5) and (6) of Table 2.1 show that using the FM regressions yields a substantial reduction in the estimated coefficients. Comparing Column (5) with Column (2), we find that in the original MW sample period, the FM regression coefficients are 98% to 127% smaller than the OLS regression coefficients across forecasting horizons. Moreover, the hypothesis that $\gamma = 0.5$ can be rejected at the 5% level across all forecasting horizons, whereas the hypothesis that $\gamma = 0$ cannot be rejected at any conventional level. Furthermore, when we extend the sample period to the end of 2020 in Column (6), we find that the FM regression coefficients are 152% to 185% smaller than the OLS regression coefficients across forecasting horizons. The hypothesis that $\gamma = 0.5$ can now be rejected at the 1% level for all forecasting horizons, whereas the hypothesis that $\gamma = 0$ cannot be rejected at any conventional level.

Columns (5) and (6) of Table 2.2 show a similarly large reduction in the regression coefficients when we switch from the OLS to the WLS regression. Comparing Column (5) with Column (2), in the original MW sample period, the WLS regression coefficients are 64% to 88% smaller than the OLS coefficients for the forecasting horizon of 3, 6, or 12 months, and 9% smaller for the forecasting horizon of one month. Comparing Column (6) with Column (3), in the extended sample period, the WLS regression coefficients are 70% to 88% smaller than the OLS coefficients for the forecasting horizon of 3, 6, or 12 months, and 17% smaller for the forecasting horizon of one month. Across all forecasting horizons, neither the hypothesis of $\gamma = 0.5$ nor the hypothesis of $\gamma = 0$ can be rejected at any conventional level in both the original MW sample period and the extended sample period.

These results suggest that the estimated regression coefficients are very sensitive to the regression methodology. So why are the FM and WLS regression coefficients so much smaller compared to the OLS regression coefficients? In Appendix B, we show that while the FM regression gives equal weights to each cross-section, the pooled OLS regression gives more weight to observations in cross-sections with larger cross-sectional variations. Intuitively, consider September 2008, the month when Lehman Brothers failed. The regression predictor $\text{SVIX}^2_{i,t,t+T}$ has a substantially higher cross-sectional variation in that month. Under the FM regression approach, the observations in this month will only affect the estimated coefficient for the September 2008 cross-section and thus affect the final FM coefficient equally as the observations in the other cross-sections. In contrast, in the pooled OLS panel regression, the observations in the September 2008 cross-section will affect the final OLS regression coefficient more because the regression predictor in September 2008 has a larger cross-sectional variance. Thus, the fact that the pooled OLS regression coefficients overweight the observations in volatile months combined with our findings that the OLS coefficients are much more positive than the FM regression coefficients suggest that the predictive relation between $\text{SVIX}^2_{i,t,t+T}$ and future stock returns is much more positive when the market is more volatile. This conjecture can also explain why the WLS regression coefficients are smaller than the OLS regression coefficients, as the former down-weight the observations with large residuals, which tend to occur when the market is more volatile.

Columns (4) and (7) of Tables 2.1 and 2.2 directly test this conjecture by using the same specification in Columns (3) and Column (6), respectively, with the only change of excluding 5% of months with the largest cross-sectional standard deviation of $\mathrm{SVIX}_{i,t,t+T}^2$ from the regression. We find that dropping these months greatly reduces the estimated regression coefficients, regardless of the regression methodology. Column (4) of Panel A of Table 2.1 finds an OLS coefficient of -0.052, 115% smaller than that in Column (3). Across forecasting horizons in Panels A–D, the OLS coefficient reduction from Column (3) to (4) ranges between 100% to 159% in Table 2.1 without firm fixed effects and between 34% to 79% in Table 2 with firm fixed effects. Comparing Column (7) with Column (6) in Table 2.1, we see that the FM coefficient drops by 19% to 39% across forecasting horizons. Comparing Column (7) with Column (6) in Table 2.2, the WLS coefficient drops by an even larger magnitude, between 193% to 296% across forecasting horizons. Therefore, dropping the months with the highest cross-sectional standard deviation of $\mathrm{SVIX}_{i,t,t+T}^2$ drastically attenuates the coefficients, rendering all of the predictive regression coefficients in Tables 1 and 2 statistically insignificant in Columns (4) and (7).

To summarize, compared to the MW estimates, the predictive regression coefficient in Eq. (1) is reduced modestly when we correct the look-ahead bias by removing the MW data filter. The predictive regression coefficient is not sensitive to the addition of six-year worth of post-publication data. However, the predictive regression coefficient is greatly reduced and is not statistically different from zero in all cases when we use the WLS or FM regressions rather than the OLS regression. We observe a similarly large reduction in the estimated coefficients when we exclude the months with the highest cross-sectional standard deviation of $\text{SVIX}^2_{i,t,t+T}$. These results indicate that the predictive power of $\text{SVIX}^2_{i,t,t+T}$ for future stock returns concentrates in volatile times, that is, months when $\text{SVIX}^2_{i,t,t+T}$ and stock returns have a high cross-sectional standard deviation. We investigate this issue using the OOS analysis of Welch and Goyal (2008) in the next subsection.

2.2.3 Out-of-Sample Analysis

Following Welch and Goyal (2008), we compare the OOS performances of the predictions based on the MW model, the historical mean model, and a model with an uninformative prior. Expected equity returns in excess of the market return based on the MW model is $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}} = \frac{1}{2}\left(\text{SVIX}^2_{i,t,t+T} - \overline{\text{SVIX}}^2_{t,t+T}\right)$. Our uninformative prior model assumes that all stocks earn the same expected returns as the market portfolio and thus $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}} = 0$ (benchmark 1) and our historical model assumes $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}}$ is equal to the historical (grand) mean of $\frac{R_{i,t,t+T}-R_{m,t,t+T}}{R_{f,t,t+T}}$ since year 1982 (benchmark 2).

We create OOS plots in Figure 2.1 by subtracting the cumulative sum-squared error (SSE) of the predictions based on the MW model from the cumulative SSE of the benchmark models. This makes it easier to see the relative performance of the forecasting model, where an increase in a line from the plot indicates better performance of the tested model compared to the benchmark models. The plots across all horizons show a consistent pattern: in the pre-2008 period, the performance of the MW model fluctuates around that of the benchmark models (near zero relative difference in the cumulative SSE); during the second half of the 2008-2009 financial crisis,[25] the MW model shows a huge outperformance compared to the two benchmark models[26]; from the end of the 2008-2009 recession to 2019, the MW model performs on par with the benchmark models; finally, during the recent COVID-19 recession period, the MW model again outperforms the benchmark models.[27] To sum up, the analysis in Figure 2.1 confirms that the superior performance of the MW model comes mostly from the second half of the Great recession when the market is more volatile. In the post-publication period, the model's predictive performance is similar to that of the benchmark models until the COVID-19 period, when the model starts to outperform once again.

2.2.4 Subsample Analysis

We investigate whether the predictive relation documented in Martin and Wagner (2019) also exhibits substantial cross-sectional variations. To do so, for each month, we sort S&P 500 stocks equally into three subsamples (high-, middle, and low-characteristics, respectively) based on popular equity investment styles. We first start by repeating the OOS analysis from Section 2.2.3 for the characteristics-sorted subsamples. We then conduct the in-sample analysis by repeating the pooled OLS panel regressions with firm fixed effects used in Martin and Wagner

[25] The turning points in all sub-figures happen after September 2008 when Lehman filed for bankruptcy.
[26] This is more clearly visible for longer horizons.
[27] The performance during the COVID-19 recession period is not shown for the 12-month horizon because the x-axis represents the prediction date. This can easily be updated once the data beyond December 2020 becomes available in the future.

(2019) for these subsamples. Panels A–C report results for subsamples sorted on valuation ratios, Panel D for subsamples sorted on market capitalization, Panels E–F for subsamples sorted on profitability ratios, and Panel G for subsamples sorted on momentum. For brevity, we report the results for only the 12-month horizon.

The OOS plots in Figure 2.2 show a clear difference in the performance of the MW model between the high- and low-characteristics subsamples. The MW model greatly underperforms the uninformative prior and historical mean benchmarks during the dot-com bubble in the growth, profitable, large-cap, and high momentum subsamples, with the difference in OOS cumulative SSE (the benchmark minus the MW model) being negative for most of the sample period. The opposite is true for the value, unprofitable, small-cap, and low momentum subsamples, in which the MW model outperforms both benchmarks throughout the sample period, with a strong increase in the OOS cumulative SSE difference during the 2008-2009 financial crisis and the COVID-19 period.

The in-sample predictive regression coefficients from Table 3 show similarly large differences between the low- and high-characteristics sorted subsamples. Columns (1) and (2) show the OLS regression results with firm fixed effects for the forecasting horizon of 12 months for brevity. Column (1) of Panel A shows that the predictive regression coefficient is 0.467 with a t-statistic of 1.10 among low book-to-market stocks. Column (2) shows that the coefficient for high book-to-market stocks is about two times as large at 0.858 with a t-statistic of 4.09. We observe similarly large differences in the estimated regression coefficients for the subsamples sorted by the other characteristics in Panels B-G. Across the board, the estimated regression coefficients among value, unprofitable, small-cap, and low momentum stocks are about two times to six times larger than the coefficients among stocks with the opposite characteristics. For the value, unprofitable, and low momentum subsamples, the hypothesis that $\gamma = 0.5$ has a p-value ranging from 0.033 to 0.119, whereas the hypothesis that $\gamma = 0$ can be rejected at the 1% level across the board. For the growth, profitable, large-cap, and high momentum subsamples, the estimated coefficients are between zero and 0.5, and neither the hypothesis that $\gamma = 0.5$ nor the hypothesis that $\gamma = 0$ can be rejected at any conventional level. Similar to the findings in Section 2.2.2, Columns (3) and (4) of Table 2.3 show that the predictive regression coefficients are substantially lower when we exclude 5% of the months with the largest cross-sectional standard deviation of $\text{SVIX}^2_{i,t,t+T}$.

Overall, both the OOS and in-sample results based on various stock characteristics-sorted subsamples indicate that the predictive relation between risk-neutral variance and future stock returns has substantial cross-sectional variations: it is much stronger among value, unprofitable, small-cap, and low momentum stocks than growth, profitable, large-cap, and high momentum stocks. Furthermore, the OOS plots for these subsamples reveal similar trends to Figure 2.1, indicating that the MW model's better performance comes from the second half of the financial crisis and the COVID-19 period.

2.2.5 Possible Explanations for Time-varying Performance

Finally, we investigate the possible reasons why the MW model performs better in some periods and worse in other periods. Figure 2.3 plots the monthly time-series of the cross-sectional 5%, 25%, 75%, and 95% percentiles of $\mathrm{SVIX}^2_{i,t,t+T}$ for the forecasting horizon of 12 months. We overlay onto this figure the cumulative sum of the squared error difference between the historical mean model (benchmark 2) and the MW model for the forecasting horizon of 12 months from Figure 2.1, with the value indicated on the right y-axis. The arrows in Figure 2.3 highlight that the periods when the MW model outperformed correspond to the periods when there are large variations in the cross-sectional variation of $\mathrm{SVIX}^2_{i,t,t+T}$.

A simple time-series regression of the squared error difference between the benchmark model and the MW model ($\mathrm{SSE_{bench}} - \mathrm{SSE_{MW}}$) on the change in the cross-sectional dispersion of $\mathrm{SVIX}^2_{i,t,t+T}$ (ChgDisp=95% percentile-5% percentile) is reported as follows

$$\mathrm{SSE_{bench,t}} - \mathrm{SSE_{MW,t}} = \underset{(-2.16)}{-0.003} + \underset{(3.99)}{0.19} \times \mathrm{ChgDisp_t} + \epsilon_t \tag{2}$$

The estimated coefficient on ChgDisp is 0.19, positive and statistically significant, with a Newey and West (1987) *t*-statistic of 3.99. The regression R^2 is very high, at 17.13%. This result suggests that one reason why the MW outperforms the benchmark model is that in periods when risks change drastically, $\mathrm{SVIX}^2_{i,t,t+T}$ may be a more timely measure than the historical risk measures in capturing real-time fluctuation in expected risk and thus expected returns.

However, why did the MW model outperform the benchmark model over the financial crisis and COVID-19 period but not the dot-com bubble period? We explore two possible reasons. First, one key assumption required to obtain the MW model in theory is by linearizing $\beta^2_{i,t} \approx 2\beta_{i,t} - 1$, which is appropriate if $\beta_{i,t}$ is sufficiently close to one. To test the accuracy of this linear approximation, we investigate whether the underperforming periods correspond to a period of large approximation errors for the term $\beta_{i,t}^2$ in the derivation of the MW model. Panel A of Figure 2.4 shows the time-series of the absolute value of the cross-sectional mean of the approximation error, that is, $|\beta_{i,t}^2 - (2\beta_{i,t} - 1)|$ across all horizons. Panel A shows that the approximation error is the largest during the dot-com bubble during the entire sample period. Panel B of Figure 2.4 plots the time-series of the percentiles of the option-implied stock betas. It shows that the deviation from the linearization point ($\beta_{i,t} = 1$) is the greatest during the dot-com bubble. Both Panels A and B suggest that the linear approximation errors can be a reason for the relative underperformance of the MW model during the dot-com bubble period.

Second, our results in Section 2.2.4 show that the predictive power of the risk-neutral variance for future stock returns is several times larger among value, unprofitable, small-cap, and past-loser stocks than among growth, profitable, large-cap, and past-winner stocks. Therefore, in periods of high volatility, the performance of MW's model may also depend on which type of stocks are more volatile in that period. Figure 4 of Martin and Wagner (2019) shows that small-cap, past-loser, and value stocks were much more volatile than large-cap, past-winner, and growth stocks during the financial crisis, but the opposite is true during the dot-com bubble period. In other words, stocks for which $\mathrm{SVIX}_{i,t,t+T}^2$ has high predictive power are more volatile during the financial crisis period, whereas these stocks were less volatile during the dot-com bubble. This may be another reason why the MW model's performance is better during the financial crisis period than during the dot-com bubble period.

2.3 Conclusion

We re-examine the main finding in Martin and Wagner (2019) that ex-ante expected excess-of-market stock returns are one-half of the excess stock variances in several ways. We first correct a look-ahead bias introduced by a data filter used in the original paper. Using the same sample period and the same pooled OLS regression specification, removing the data filter reduces the estimated predictive coefficient in Eq. (1) by about 16% to 18% and 26% to 34% across forecasting horizons of 1, 3, 6, and 12 months for the specifications with and without firm fixed effects, respectively. We then extend the sample by six years to the end of 2020 and find that the magnitude of the estimated regression coefficient is not sensitive to the sample extension.

We find that the estimated predictive coefficient is sensitive to the regression methodology due to the time variations in the predictive relation. When we replace the pooled OLS regressions used in the original paper with the FM regressions, we find that the estimated coefficients shrink between 98% to 127% in the original sample period and 152% to 185% in the extended sample period across forecasting horizons. In the extended sample period, the hypothesis that the predictive coefficient in Eq. (1) is equal to 0.5 can be rejected at the 1% level across all forecasting horizons. At the same time, the estimated coefficients are no longer statistically significantly different from zero across all forecasting horizons. We also find a large reduction in the estimated coefficients when we replace the OLS regressions with the WLS regressions. The OLS regression coefficients are substantially higher than the FM and WLS coefficients because 1) the OLS estimates are more influenced by observations in high-volatility months, and 2) the predictive relation is stronger in these high-volatility months. Consistent with this intuition, when we exclude 5% of months with the highest cross-sectional standard deviation of the predictor, we again find that the estimated regression coefficients become insignificantly different from zero for all forecasting horizons.

We also find that the predictive coefficient also exhibits large cross-sectional variations, with the OLS estimates being two to six times larger for value, unprofitable, small-cap, and past-loser stocks than for growth, profitable, large-cap, and past-winner stocks. Consistent with our in-sample analysis, the out-of-sample outperformance of the MW model comes almost entirely from the high volatility periods – the financial crisis and the COVID-19 periods, with the exception of the dot-com bubble period. We provide tentative explanations of the time-varying performance of the MW model. Altogether, our results reveal substantial time-series and cross-sectional variations in the predictive relation between risk-neutral variance and future stock returns.

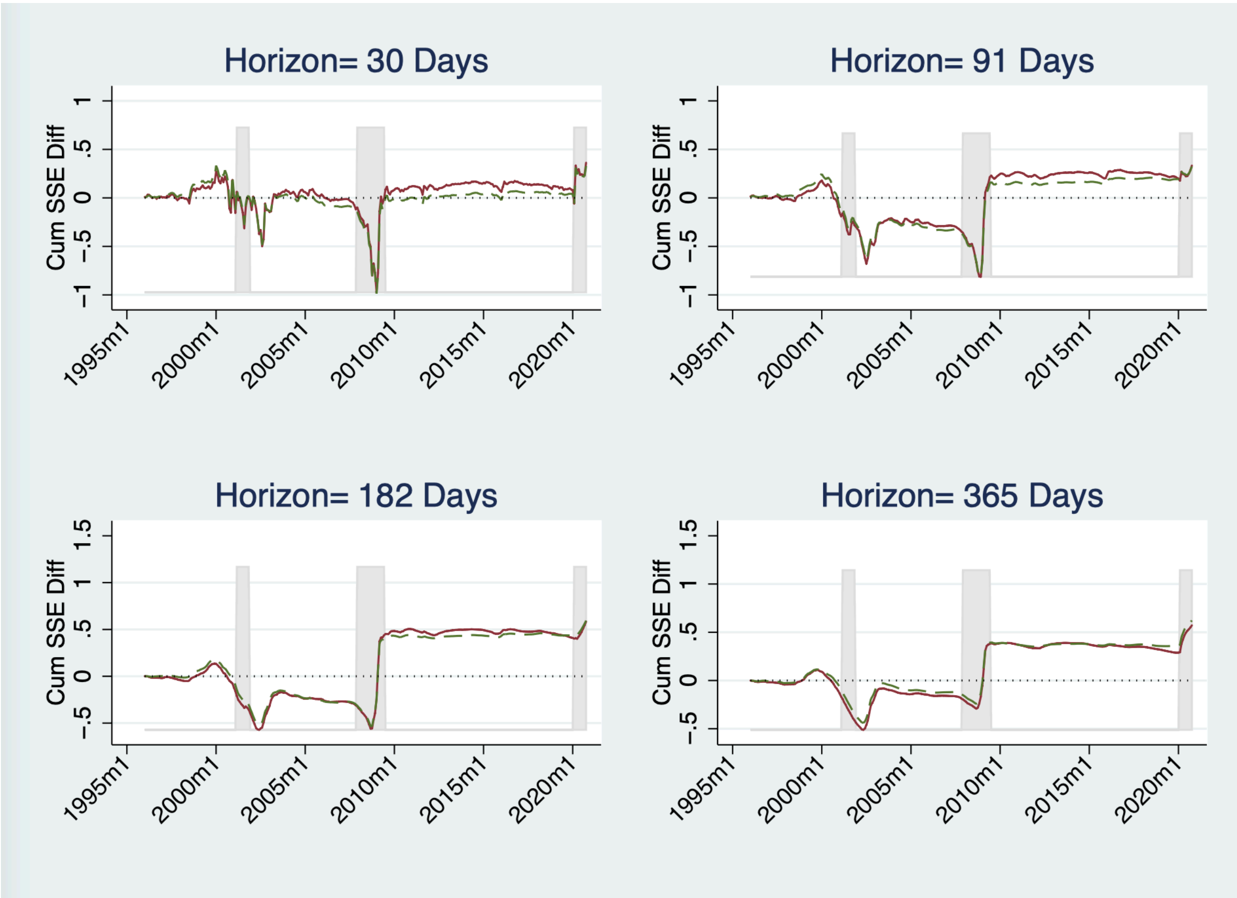
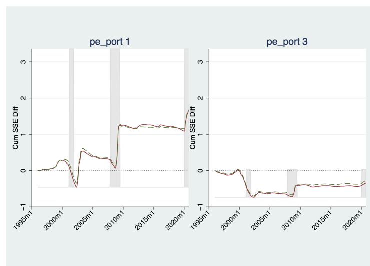


Figure 2.1: OOS Cumulative SSE Difference

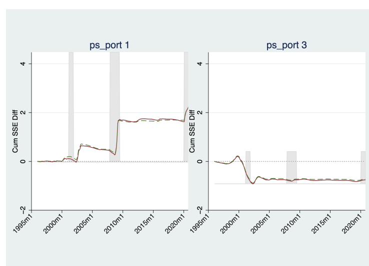
Description: This figure plots the OOS performance following Welch and Goyal (2008) by subtracting the cumulative sum-squared error (SSE) of the predictions based on the MW model from the cumulative SSE of the benchmark models for the forecasting period from January 1996 to December 2020. The model used here is $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}} = \frac{1}{2}\left(\mathrm{SVIX}^2_{i,t,t+T} - \overline{\mathrm{SVIX}}^2_{t,t+T}\right)$ from Martin and Wagner (2019). The solid line shows the results when using $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}} = 0$ as the benchmark and the dash line shows the results when the historical (grand) mean of $\frac{E_t(R_{i,t,t+T}-R_{m,t,t+T})}{R_{f,t,t+T}}$ since year 1982 is used as the benchmark. An increase (decrease) in a line from the plot indicates better (worse) performance of the tested model compared to the benchmark. The official NBER recession periods are used for the dot-com bubble, 2008-2009 financial crisis, and COVID-19, shaded in grey.

Panel A - Low Book-to-Market vs. High Book-to-Market

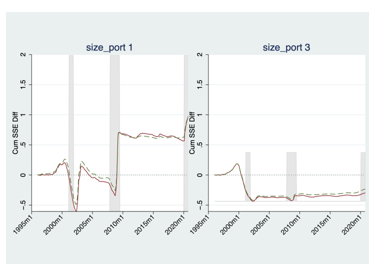
Panel B - Low Price-to-Earnings vs. High Price-to-Earnings



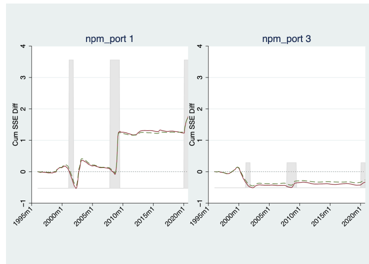
Panel C - Low Price-to-Sales vs. High Price-to-Sales



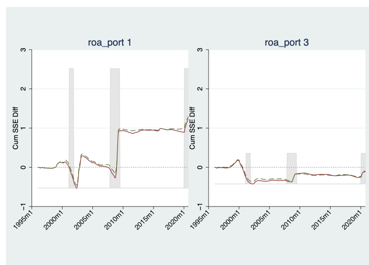
Panel D - Small-Cap vs. Large-Cap



Panel E - Low Net Profit Margin vs. High Net Profit Margin



Panel F - Low Return on Assets vs. High Return on Assets



Panel G - Low Momentum vs. High Momentum

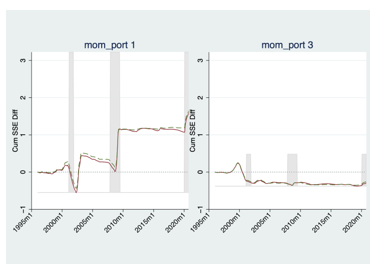


Figure 2.2: Subsample Analysis of OOS Cumulative SSE Difference

Description: See Figure 2.1, except now we examine the MW model's performance by sorting S&P 500 stocks equally into three subsamples (high-, middle, and low-characteristics, respectively) based on popular equity investment styles for each month. We report the results for only the low-characteristics (port 1) and high- characteristics (port 3) subsamples for the 12-month horizon for brevity. The same scale is used for both plots within each panel.

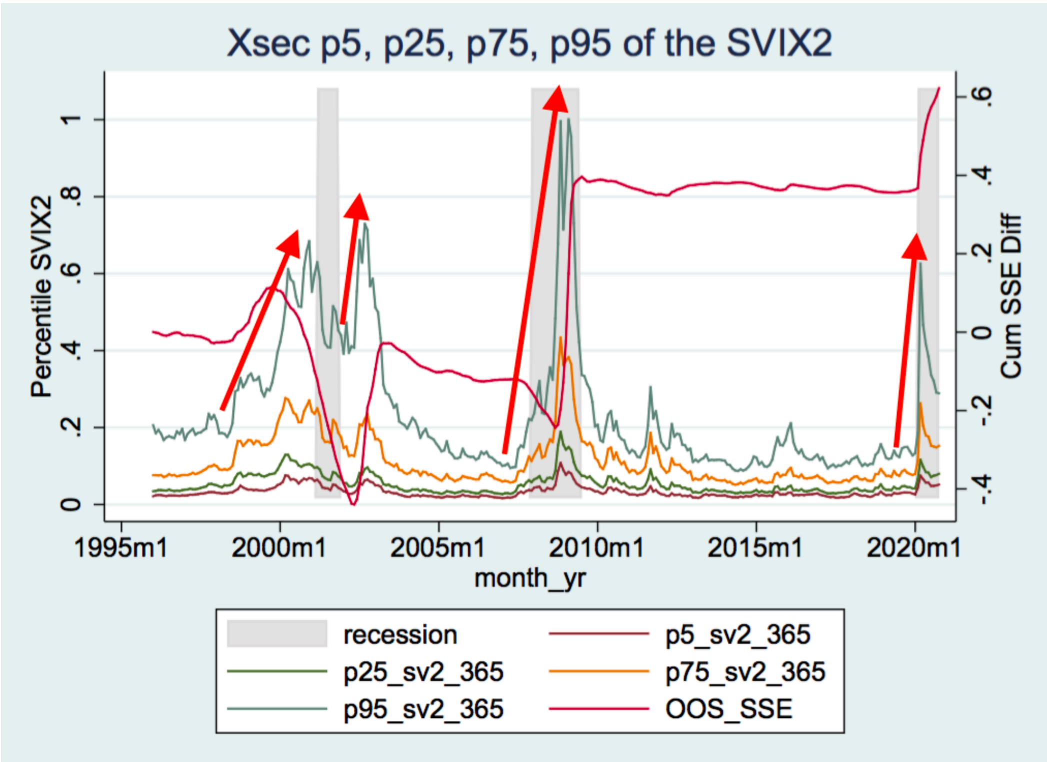
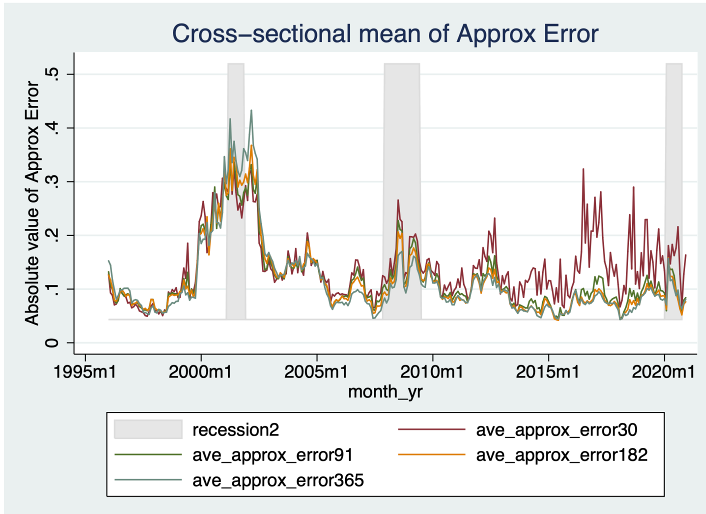


Figure 2.3: Time-series of the Cross-sectional Percentiles of $\mathrm{SVIX}^2_{i,t,t+T}$

Description: This figure plots the monthly time-series of the cross-sectional 5%, 25%, 75%, and 95% percentiles of $\mathrm{SVIX}^2_{i,t,t+T}$ for the forecasting horizon of 12 months. The y-axis on the right is for the cumulative sum of the difference between the squared predictions errors between the historical mean model (benchmark 2) and the MW model for the forecasting horizon of 12 months from Figure 2.1. The arrows represent periods when there are large variations in the cross-sectional dispersion of $\mathrm{SVIX}^2_{i,t,t+T}$.

Panel A - Time-series of the Cross-sectional Mean of the Approximation Error



Panel B - Time-series of the Percentiles of Option-Implied Beta

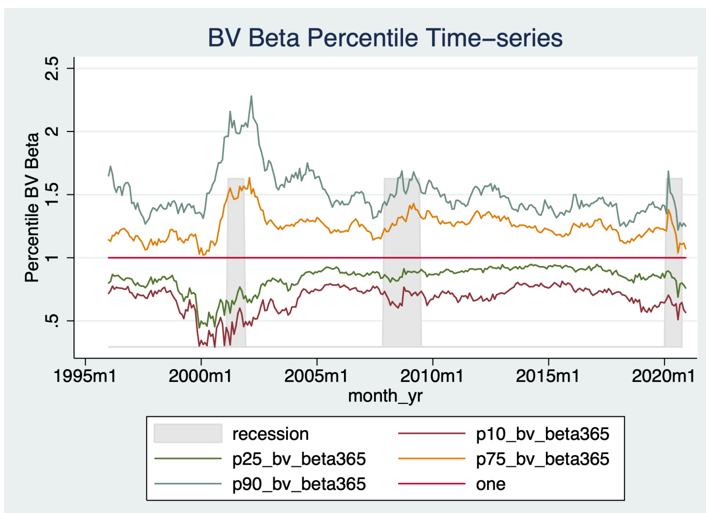


Figure 2.4: Beta Analysis

Description: This figure conducts beta analysis using option-implied stock betas. Panel A shows the time-series of the absolute value of the cross-sectional mean of the approximation error $|\beta_{i,t}^2 - (2\beta_{i,t} - 1)|$ across forecasting horizons of 1, 3, 6, and 12 months. Panel B plots the time-series of the percentiles of the option-implied stock betas for the forecasting horizon of 12 months.

Table 2.1: Forecasting Stock Returns using Pooled Panel Regressions and FM Regressions

Description: This table presents results from regressing stock returns in excess of the market onto risk-neutral excess stock variance for S&P 500 firms. The data is monthly, and the return horizons match the maturities of the options used to compute the risk-neutral variance. Panels A to D are for return horizons equal to 12, 6, 3, and 1 month, respectively. Column (1) reports panel regression results using the original sample period in Martin and Wagner (2019) from January 1996 to October 2014, and Column (2) removes the MW data filter. "MW data filter" refers to the exclusion of stock-month observations for S&P 500 stocks 1, 3, 6, 12 months before a stock is deleted from the S&P 500 index for forecasting horizons 1, 3, 6, and 12 months, respectively. Column (3) extends the sample period from Column (2) to the end of 2020. Columns (5) and (6) use the same specifications in Columns (2) and (3) with the only change of using the Fama and MacBeth (1973) (FM) regression. Columns (4) and (7) use the same specifications in Columns (3) and (6) with the only change of excluding 5% of months with the largest cross-sectional standard deviation in $\text{SVIX}^2_{i,t,t+T}$. Standard errors obtained from the block bootstrap procedure described in MW are reported in parentheses. In each panel, we report p-values of Wald tests of the null hypotheses.

Panel A - 12-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	0.016	0.017	0.005	0.006	0.021	0.009	0.006
	(0.019)	(0.018)	(0.014)	(0.014)	(0.017)	(0.013)	(0.014)
γ	0.536	0.352	0.338	-0.052	0.008	-0.191	-0.245
	(0.294)	(0.270)	(0.263)	(0.196)	(0.207)	(0.201)	(0.194)
$H_0 : \gamma = 0.5$	0.902	0.583	0.537	0.005	0.018	0.001	0.000
$H_0 : \gamma = 0$	0.068	0.192	0.199	0.789	0.971	0.342	0.207
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	FM	FM	FM

Panel B - 6-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	0.015	0.016	0.005	0.006	0.021	0.010	0.007
	(0.016)	(0.016)	(0.013)	(0.013)	(0.015)	(0.012)	(0.012)
γ	0.526	0.381	0.375	-0.065	-0.034	-0.251	-0.312
	(0.317)	(0.307)	(0.291)	(0.227)	(0.223)	(0.204)	(0.214)
$H_0 : \gamma = 0.5$	0.935	0.699	0.667	0.013	0.017	0.000	0.000
$H_0 : \gamma = 0$	0.097	0.214	0.198	0.776	0.878	0.218	0.144
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	FM	FM	FM

Panel C - 3-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	0.017	0.019	0.007	0.007	0.022	0.012	0.009
	(0.015)	(0.014)	(0.012)	(0.012)	(0.014)	(0.011)	(0.011)
γ	0.401	0.297	0.312	0.001	-0.040	-0.265	-0.315
	(0.290)	(0.286)	(0.258)	(0.218)	(0.220)	(0.198)	(0.216)
$H_0 : \gamma = 0.5$	0.732	0.478	0.467	0.022	0.014	0.000	0.000
$H_0 : \gamma = 0$	0.168	0.299	0.226	0.996	0.855	0.182	0.144
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	FM	FM	FM

Panel D - 1-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	0.016	0.018	0.007	0.013	0.022	0.012	0.010
	(0.014)	(0.015)	(0.012)	(0.012)	(0.014)	(0.011)	(0.011)
γ	0.363	0.263	0.316	-0.187	-0.071	-0.163	-0.227
	(0.307)	(0.303)	(0.288)	(0.300)	(0.233)	(0.206)	(0.216)
$H_0 : \gamma = 0.5$	0.656	0.433	0.523	0.022	0.014	0.001	0.001
$H_0 : \gamma = 0$	0.238	0.386	0.272	0.533	0.760	0.428	0.294
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	FM	FM	FM

Table 2.2: Forecasting Stock Returns using Panel Regressions and Weighted-Least-Squares Regressions with Fixed Effects

Description: See Table 2.1, except now firm fixed effects are included, and Columns (5) to (7) report results using weighted-least-squares (WLS) regressions. We model the heteroscedasticity by regressing the absolute value of the residuals from Columns (2) to (4) on risk-neutral excess stock variance. We then use the inverse of the fitted values as the weights in the WLS regressions in Columns (5) to (7).

Panel A - 12-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\sum_i w_i \alpha_i$	0.035	0.034	0.026	0.026	0.036	0.028	0.027
	(0.008)	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
γ	0.873	0.714	0.712	0.292	0.197	0.123	-0.162
	(0.278)	(0.253)	(0.261)	(0.230)	(0.253)	(0.497)	(0.232)
$H_0 : \gamma = 0.5$	0.179	0.398	0.416	0.365	0.232	0.447	0.004
$H_0 : \gamma = 0$	0.002	0.005	0.006	0.204	0.437	0.805	0.485
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	WLS	WLS	WLS

Panel B - 6-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\sum_i w_i \alpha_i$	0.034	0.033	0.024	0.024	0.035	0.026	0.025
	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.009)	(0.006)
γ	0.936	0.767	0.768	0.320	0.275	0.233	-0.223
	(0.350)	(0.329)	(0.314)	(0.239)	(0.397)	(1.029)	(0.243)
$H_0 : \gamma = 0.5$	0.213	0.417	0.393	0.452	0.571	0.795	0.003
$H_0 : \gamma = 0$	0.007	0.020	0.014	0.181	0.488	0.821	0.358
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	WLS	WLS	WLS

Panel C - 3-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\sum_i w_i \alpha_i$	0.035	0.034	0.025	0.024	0.034	0.026	0.025
	(0.006)	(0.006)	(0.006)	(0.006)	(0.020)	(0.018)	(0.011)
γ	0.754	0.636	0.665	0.391	0.078	0.079	-0.155
	(0.325)	(0.320)	(0.288)	(0.228)	(2.137)	(2.789)	(1.686)
$H_0 : \gamma = 0.5$	0.434	0.670	0.567	0.633	0.843	0.880	0.698
$H_0 : \gamma = 0$	0.020	0.047	0.021	0.086	0.971	0.977	0.927
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	WLS	WLS	WLS

Panel D - 1-month horizon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\sum_i w_i \alpha_i$	0.036	0.036	0.027	0.027	0.036	0.028	0.028
	(0.007)	(0.007)	(0.006)	(0.006)	(0.020)	(0.035)	(0.008)
γ	0.676	0.554	0.617	0.129	0.505	0.510	-0.474
	(0.336)	(0.331)	(0.299)	(0.318)	(2.781)	(4.324)	(1.260)
$H_0 : \gamma = 0.5$	0.600	0.871	0.695	0.243	0.999	0.998	0.439
$H_0 : \gamma = 0$	0.044	0.094	0.039	0.686	0.856	0.906	0.707
MW Data Filter	Y	N	N	N	N	N	N
Sample Period	MW	MW	Ex	Ex	MW	Ex	Ex
Regression	OLS	OLS	OLS	OLS	WLS	WLS	WLS

Table 2.3: Subsample Analysis - Forecasting Stock Returns using Panel Regressions with Fixed Effects

Description: This table presents panel regression results from regressing stock returns in excess of the market onto risk-neutral excess stock variance for S&P 500 firms after including firm fixed effects. We sort S&P 500 stocks equally into three subsamples (high-, middle, and low-characteristics, respectively) based on popular equity investment styles (Panels A–G) for each month. We report the results using the extended sample period from January 1996 to December 2020 for only the low-characteristics (Column 1) and high- characteristics (Column 2) subsamples for the 12-month horizon for brevity. Columns (3) and (4) use the same specifications in Columns (1) and (2) with the only change of excluding 5% of months with the largest cross-sectional standard deviation in $\mathrm{SVIX}_{i,t,t+T}^2$. In each panel, we report the p-values of Wald tests of the null hypotheses.

Panel A - Low Book-to-Market vs. High Book-to-Market

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.047	0.003	0.048	-0.001
	(0.006)	(0.004)	(0.007)	(0.004)
γ	0.467	0.858	0.122	0.554
	(0.424)	(0.210)	(0.357)	(0.222)
$H_0 : \gamma = 0.5$	0.939	0.087	0.290	0.808
$H_0 : \gamma = 0$	0.270	0.000	0.733	0.013

Panel B - Low Price-to-Earnings vs. High Price-to-Earnings

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.049	0.014	0.046	0.014
	(0.006)	(0.006)	(0.007)	(0.006)
γ	0.874	0.380	0.464	0.034
	(0.240)	(0.302)	(0.225)	(0.246)
$H_0 : \gamma = 0.5$	0.119	0.692	0.874	0.058
$H_0 : \gamma = 0$	0.000	0.208	0.039	0.891

Panel C - Low Price-to-Sales vs. High Price-to-Sales

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.022	0.032	0.018	0.032
	(0.004)	(0.005)	(0.003)	(0.005)
γ	0.985	0.182	0.592	0.002
	(0.228)	(0.333)	(0.197)	(0.304)
$H_0 : \gamma = 0.5$	0.033	0.339	0.640	0.101
$H_0 : \gamma = 0$	0.000	0.584	0.003	0.994

Panel D - Small-Cap vs. Large-Cap

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.042	0.018	0.026	0.019
	(0.003)	(0.005)	(0.003)	(0.005)
γ	0.748	0.131	0.379	-0.144
	(0.235)	(0.382)	(0.222)	(0.396)
$H_0 : \gamma = 0.5$	0.292	0.333	0.586	0.104
$H_0 : \gamma = 0$	0.001	0.732	0.087	0.716

Panel E - Low Net Profit Margin vs. High Net Profit Margin

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.033	0.029	0.030	0.029
	(0.006)	(0.005)	(0.006)	(0.006)
γ	0.970	0.195	0.545	-0.047
	(0.249)	(0.326)	(0.240)	(0.333)
$H_0 : \gamma = 0.5$	0.059	0.348	0.850	0.101
$H_0 : \gamma = 0$	0.000	0.550	0.023	0.889

Panel F - Low Return on Assets vs. High Return on Assets

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.001	0.048	-0.001	0.049
	(0.005)	(0.007)	(0.005)	(0.008)
γ	0.851	0.325	0.478	0.024
	(0.222)	(0.438)	(0.225)	(0.424)
$H_0 : \gamma = 0.5$	0.114	0.690	0.924	0.262
$H_0 : \gamma = 0$	0.000	0.457	0.034	0.954

Panel G - Low Momentum vs. High Momentum

	(1)	(2)	(3)	(4)
$\sum_i w_i \alpha_i$	0.013	0.044	0.006	0.049
	(0.005)	(0.008)	(0.004)	(0.008)
γ	0.878	0.219	0.505	0.114
	(0.233)	(0.332)	(0.224)	(0.303)
$H_0 : \gamma = 0.5$	0.105	0.397	0.983	0.203
$H_0 : \gamma = 0$	0.000	0.509	0.024	0.706

2.4 Appendix

Appendix A: Data and Replication Procedure

We follow Martin and Wagner (2019) in constructing the stock-level $\text{SVIX}^2_{i,t,t+T}$. For each stock i, $\Omega_{i,t,t+T}(K)$ is the time t price of an out-of-the-money option with strike K and maturity $t+T$:

$$\Omega_{i,t,t+T}(K) \equiv \begin{cases} \text{put}_{i,t,t+T}(K) & \text{if } K < F_{i,t,t+T} \\ \text{call}_{i,t,t+T}(K) & \text{if } K \geq F_{i,t,t+T} \end{cases};$$

$\text{SVIX}^2_{i,t,t+T}$ is the risk-neutral variance for stock returns between t and $t+T$. It is theoretically equal to the integral $\int_0^\infty \Omega_{i,t,t+T}(K)dK$, which is approximated by using $\Omega_{i,t,t+T}(K)$ at discrete strike prices:

$$\text{SVIX}^2_{i,t,t+T} = \frac{2}{R_{f,t,t+T}S_t^2}\sum_{j=1}^{N}\Omega_{i,t,t+T}(K_j)\triangle K_j,$$

where $K_1, ..., K_N$ are the strikes of observable options and $\triangle K_j = \frac{K_{j+1}-K_{j-1}}{2}$, and $R_{f,t,t+T}$ is the annualized risk-free rate for the horizon T. Following Martin and Wagner (2019), we examine four horizons, T =1 month, 3 months, 6 months, and 12 months. The return forecasting horizon is matched with the option maturity, that is, we forecast 1-month stock returns using $\text{SVIX}^2_{i,t,t+T}$ based on options with 1-month maturity, 3-month stocks return using 3-month options, so on, and so forth. All options data are from OptionMetrics. The option prices are from the volatility surface file, the corresponding forward price is from the standardized options file, and the risk-free rate is from the zero-coupon yield curve. The current release of the volatility surface file contains 17 OTM options, while the previous release used by Martin and Wagner (2019) contains only 13 OTM options.[6] We thus construct two versions of $\text{SVIX}^2_{i,t,t+T}$: one uses the full 17 OTM options provided in the current release, and the other version drops the two OTM puts (calls) with the lowest (highest) strike prices. We use the latter version in our main specification for ease of comparison to the results in Martin and Wagner (2019).

Appendix B: FM Regression vs. Pooled OLS Panel Regression

We use i, t to denote firm i and month t. The predictive regression has the following form:

$$y_{i,t} = a + bx_{i,t} + \epsilon_{i,t}$$

Since we study S&P 500 firms and thus the number of firms is roughly stable over time, to simplify the notation, we assume the number of observations per cross-section N_t is a constant N. Then the estimates of the pooled OLS regression coefficients and the FM regression coefficients are as follows:

$$\begin{aligned}
\hat{b}_{\text{Pooled OLS}} &= \frac{\sum_{i,t}\left(x_{i,t} - \frac{\sum_{i,t} x_{i,t}}{T\times N}\right)\left(y_{i,t} - \frac{\sum_{i,t} y_{i,t}}{T\times N}\right)}{\sum_{i,t}\left(x_{i,t} - \frac{\sum_{i,t} x_{i,t}}{T\times N}\right)^2} \\
&= \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}\left(x_{i,t} - \frac{\sum_{i,t} x_{i,t}}{T\times N}\right)\left(y_{i,t} - \frac{\sum_{i,t} y_{i,t}}{T\times N}\right)}{\sum_{t=1}^{T}\sum_{i=1}^{N}\left(x_{i,t} - \frac{\sum_{i,t} x_{i,t}}{T\times N}\right)^2} \\
\hat{b}_{\text{FM}} &= \sum_{t=1}^{T}\frac{1}{T}\times\frac{\sum_{i=1}^{N}\left(x_{i,t} - \frac{\sum_{i} x_{i,t}}{N}\right)\left(y_{i,t} - \frac{\sum_{i} y_{i,t}}{N}\right)}{\sum_{i=1}^{N}\left(x_{i,t} - \frac{\sum_{i} x_{i,t}}{N}\right)^2}
\end{aligned}$$

To simplify the notation, let's denote $\overline{x} = \frac{\sum_{i,t} x_{i,t}}{T\times N}$, $\overline{y} = \frac{\sum_{i,t} y_{i,t}}{T\times N}$, $\overline{x}_t = \frac{\sum_{i=1}^{N} x_{i,t}}{N}$,

[6] We thank the authors for providing this information.

$V_{x,t}^{CS} = \frac{\sum_{i=1}^{N}(x_{i,t}-\overline{x}_t)^2}{N}$, $\frac{\sum_{i=1}^{N}(x_{i,t}-\overline{x})^2}{N} = \frac{\sum_{i=1}^{N}(x_{i,t}-\overline{x}_t+\overline{x}_t-\overline{x})^2}{N} = V_{x,t}^{CS} + (\overline{x}_t - \overline{x})^2$,
and $C_{x,y,t}^{CS} = \frac{\sum_{i=1}^{N}\left(x_{i,t}-\frac{\sum_i x_{i,t}}{N}\right)\left(y_{i,t}-\frac{\sum_i y_{i,t}}{N}\right)}{N}$

$$\hat{b}_{\text{FM}} = \sum_{t=1}^{T} \frac{1}{T} \times \frac{C_{x,y,t}^{CS}}{V_{x,t}^{CS}} \tag{3}$$

$$\begin{aligned}
\hat{b}_{\text{Pooled OLS}} &= \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}\left(x_{i,t}-\overline{x}_t+\overline{x}_t-\overline{x}\right)\left(y_{i,t}-\overline{y}_t+\overline{y}_t-\overline{y}\right)}{\sum_{t=1}^{T} N\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)} \\
&= \frac{\sum_{t=1}^{T}\left[N \times C_{x,y,t}^{CS} + N\left(\overline{x}_t-\overline{x}\right)\left(\overline{y}_t-\overline{y}\right)\right]}{\sum_{t=1}^{T} N\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)} \\
&= \frac{\sum_{t=1}^{T}\left[C_{x,y,t}^{CS} + \left(\overline{x}_t-\overline{x}\right)\left(\overline{y}_t-\overline{y}\right)\right]}{\sum_{t=1}^{T}\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)} \\
&= \sum_{t=1}^{T}\frac{1}{T} \times \frac{C_{x,y,t}^{CS}}{V_{x,t}^{CS}} \times \frac{T \times V_{x,t}^{CS}}{\sum_{t=1}^{T}\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)} + \frac{\sum_{t=1}^{T}\left[\left(\overline{x}_t-\overline{x}\right)\left(\overline{y}_t-\overline{y}\right)\right]}{\sum_{t=1}^{T}\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)}
\end{aligned}$$

Since $V_{x,t}^{CS}$ is much larger than $\left(\overline{x}_t - \overline{x}\right)^2$ in the data, the main intuition can be obtained by considering $\frac{(\overline{x}_t-\overline{x})^2}{V_{x,t}^{CS}} << 1$.

$$\begin{aligned}
\hat{b}_{\text{Pooled OLS}} &= \sum_{t=1}^{T}\frac{1}{T} \times \frac{C_{x,y,t}^{CS}}{V_{x,t}^{CS}} \times \frac{T \times V_{x,t}^{CS}}{\sum_{t=1}^{T} V_{x,t}^{CS}} + \frac{\sum_{t=1}^{T}\left(\overline{x}_t-\overline{x}\right)^2}{\sum_{t=1}^{T}\left(V_{x,t}^{CS}+\left(\overline{x}_t-\overline{x}\right)^2\right)} \times \frac{\sum_{t=1}^{T}\left[\left(\overline{x}_t-\overline{x}\right)\left(\overline{y}_t-\overline{y}\right)\right]}{\sum_{t=1}^{T}\left(\overline{x}_t-\overline{x}\right)^2} \\
&\sim \sum_{t=1}^{T}\frac{1}{T} \times \frac{C_{x,y,t}^{CS}}{V_{x,t}^{CS}} \times \frac{T \times V_{x,t}^{CS}}{\sum_{t=1}^{T} V_{x,t}^{CS}} \qquad (4)
\end{aligned}$$

Comparing Eqs. (3) and (4), we can see that the pooled OLS regressions give more weights to months in which $V_{x,t}^{CS}$ is relatively larger.

References

Antweiler, W., and M. Z. Frank. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards." *The Journal of Finance,* 59 (2004), 1259–1294.

Barber, B. M., and T. Odean. "Too Many Cooks Spoil the Profits: Investment Club Performance." Financial Analysts Journal, 56 (2000), 17–25

Barber, B. M., and T. Odean. "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors." *The Review of Financial Studies*, 21 (2008), 785–818.

Barber, B. M.; X. Huang; T. Odean; and C. Schwarz. "Attention Induced Trading and Returns: Evidence from Robinhood Users." Working Paper (2020).

Bartov, E.; L. Faurel; and P. S. Mohanram. "Can Twitter Help Predict Firm-Level Earnings and Stock Returns?" *The Accounting Review,* 93 (2018), 25–57.

Biggerstaff, L; D. Cicero; and M. B. Wintoki. "Insider trading patterns" *Journal of Corporate Finance,* 64 (2020).

Blackburne, T.; J. D. Kepler; P. J. Quinn; and D. Taylor. "Undisclosed SEC Investigations." *Management Science*, (2020).

Boehmer, E.; C. M. Jones; X. Zhang; and X. Zhang. "Tracking Retail Investor Activity." The Journal of Finance, 76 (2021), 2249–2305.

Chen, H.; P. De; Y. J. Hu; and B. H. Hwang. "Wisdom of Crowds: The Value of Stock Opinions Transmitted through Social Media." *Review of Financial Studies*, 27 (2014), 1367–403.

Cookson, J. A.; J. Engelberg; and W. Mullins. "Echo chambers." Working paper, (2020).

Datta, S., and M. E. Datta. "Does insider trading have information content for the bond market?" *Journal of Banking and Finance,* 20 (1996), 555–575.

Dewally, M. "Internet Investment Advice: Investing with a Rock of Salt." *Financial Analysts Journal*, 59 (2003), 65–77.

Eaton, G. W.; T. C. Green; B. S. Roseman; and Y. Wu. "Zero-Commission Individual Investors, High Frequency Traders, and Stock Market Quality." Working paper (2021).

Fama, E. F., and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Fama, E. F., and K. R. French. "A five-factor asset pricing model." *Journal of Financial Economics*. 116 (2015), 1–22.

Fama, E. F., and K. R. French. "Choosing Factors." *Journal of Financial Economics,* 128 (2018), 234–252.

Fama, E. F., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy,* 81 (1973), 607–36.

Farrell, M.; T. C. Green; R. Jame; and S. Markov. "The Democratization of Investment Research and the Informativeness of Retail Investor Trading." *Journal of Financial Economics,* (2021).

Giannini, R.; P. Irvine; and T. Shu. "Nonlocal Disadvantage: An Examination of Social Media Sentiment." *The Review of Asset Pricing Studies,* 8 (2018), 293–336.

Heimer, R. Z. "Peer Pressure: Social Interaction and the Disposition Effect." *The Review of Financial Studies*, 29 (2016), 3177–3209.

Hutto, C., and E. Gilbert. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. 8 (2014), No. 1.

Hvidkjaer, S. "Small Trades and the Cross-Section of Stock Returns." *The Review of Financial Studies,* 21 (2008), 1123–1151.

Jegadeesh, N., and S. Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *The Journal of Finance,* 48, (1993), 65–91.

Kaniel, R.; S. Liu; G. Saar; and S. Titman. "Individual Investor Trading and Return Patterns around Earnings Announcement." *The Journal of Finance*, 67 (2012), 639–680.

Lakonishok, J., and I. Lee. "Are Insider Trades Informative?" *The Review of Financial Studies,* 14 (2001), 79–111.

Martin, I. W. R., and C. Wagner. "What is the Expected Return on a Stock?" *The Journal of Finance,* 74 (2019), 1887–1929.

Newey, W. K., and K. D. West. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55 (1987), 703–708.

Renault, T. "Intraday online investor sentiment and return patterns in the U.S. stock market" *Journal of Banking and Finance*, 84 (2017), 25–40.

Sprenger, T. O.; A. Tumasjan; P. G. Sandner; and I. M. Welpe. "Tweets and Trades: The Information Content of Stock Microblogs." *European Financial Management,* 20 (2014), 926–57.

Tetlock, P. C. "All the News That's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies,* 24 (2011), 1481–512.

Welch, I, and A. Goyal. "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction." *Review of Financial Studies,* 21 (2008), 1455–1508.

Welch, I. "The Wisdom of the Robinhood Crowd." *The Journal of Finance, Forthcoming* (2021).