ESSAYS IN FINANCIAL ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Luoyi Su

May 2019

ESSAYS IN FINANCIAL ECONOMICS

Luoyi Su, Ph.D.

Cornell University 2019

This dissertation uses novel data to provide micro-level perspectives on the behavior of fund managers and investors. Chapter 1 examines the effect of a trader's personal co-investment or skin-in-the-game on fund risk-taking. Using a unique dataset from an online social trading platform, I uncover a source of exogenous variation in trader's skin-in-the-game to investigate the causal effect of skin-in-the-game on fund risk-taking. I find that having no skin-in-the-game significantly increases trader's incentive for risk-taking. The findings provide evidence in support of skin-in-the-game as an important mechanism to align incentives of traders and investors.

Chapter 2 studies individual fund investor extrapolation. In particular, I examine fund investment and withdrawal events at individual investor level. The sample is constructed from brokerage account data of global retail forex investors from the social trading platform. The findings provide evidence of fund investor extrapolation, in which the past performance consistency of fund investment has a significant effect on investor's withdrawal decision. This effect is more pronounced for investors from more developed countries. The results also highlight that investor's withdrawal decision depends not only on past performance and volatility but also on the consistency of past performance. However, none of these factors positively predicts future performance. These results support the view that fund investors over-extrapolate as they tend to extrapolate based on past performance measures and do not profit from doing so.

As U.S. adults increasingly obtain news through mobile devices rather than desktop computers, Chapter 3 compares “mobile sentiment” with “desktop sentiment” in predicting future stock returns and liquidity. I construct unique data scraped from Google, which sometimes produces very different results on mobile vs. desktop search due to different ranking practices (e.g. a link with text consisting of negative words about a stock is shown on mobile but not on desktop) . Thus, I collect daily Google search results separately on mobile and desktop platforms for tickers of stocks in the S&P 500 index. I conduct textual analysis on the search results. I find that negative mobile or desktop sentiment predicts abnormal return reversal in the following week with mobile serving as a more significant predictor than desktop. I show that this reversal is mainly driven by stock over-pricing. That is as investors become more optimistic due to recent good news, the stock is over-priced and will later revert back to its fundamental value (i.e. lower future returns), whereas as investors become more pessimistic due to recent bad news, the stock is unlikely to be under-priced and have reversal. In addition, the effect of mobile sentiment on returns becomes more pronounced than desktop sentiment in stocks of high retail interests. This supports the idea that going mobile is a preferred way to obtain trading information among less sophisticated investors. I find weak evidence that mobile sentiment relates more to liquidity measured by effective spreads and volume than desktop sentiment. The results also suggest that sentiment is mutually Granger-causal with either return or liquidity. In the end, my results highlight the growing relevance of mobile media in disseminating financial news, and provide suggestive evidence that compared with desktop computer users, mobile users are less informed and more akin to sentiment investors.

BIOGRAPHICAL SKETCH

Luoyi Su grew up in Shanghai, China. He went to the U.S. in 2007 to study at a high school in New Jersey. The financial crisis at the time triggered his interest in economics and financial markets. He went on to develop his academic interests at Cornell University and obtained his Bachelor of Arts degree in economics and mathematics in 2014. He then continued to pursue his PhD at Cornell Economics Department studying financial economics.

To my family

ACKNOWLEDGEMENTS

I owe a deep debt of gratitude to my committee members: Nicholas M. Kiefer (Chair), Byoung-Hyoun Hwang, and Karl Shell. I am deeply indebted to Nick for his invaluable guidance and advice, as well as for his kind patience and time in supervising my studies. I deeply appreciate the many helpful suggestions from Byoung, whose powerful insights have led me to interesting research findings. I'm deeply grateful to Karl for guiding me with his wisdom to navigate me through various stages of the learning process. This dissertation would not have been possible without their wonderful guidance and help.

I sincerely thank the many professors for their generous help and advice: Panle J. Barwick, David Easley, Yongmiao Hong, Robert Jarrow, Andrew Karolyi, Peng Liu, Maureen O'Hara, Scott Yonker and others.

I am also very grateful to Jennifer Wissink for her great tips on teaching and the honor to serve as the trainer of the teaching assistants in the Department of Economics.

I also thank Cornell University for the financial support through Sage Fellowship, and the great staff in the Department of Economics.

Lastly, I thank my family and friends for their love and unwavering support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

EFFECT OF TRADER'S CO-INVESTMENT ON FUND RISK-TAKING

1.1 Introduction

With the rapid growth of the mutual fund and hedge fund industries over the past decade, the SEC has increased regulatory oversight and disclosure requirements to better protect investors. Specifically, the SEC requires mutual funds to disclose fund management's ownership annually in the funds' Statement of Additional Information starting March 2005. Then in 2011, the Dodd-Frank Act required hedge funds with more than $150 million to register with the SEC and file a Form ADV, in which questions about the insider investment of the funds were added in 2012[1]. These regulations aimed at improving transparency and oversight have brought attention to the role of fund managers' incentives on risk-taking. One proxy for managerial incentives is managers' ownership stakes or co-investment in their funds (hereafter skin-in-the-game). In this paper, I examine how skin-in-the-game affects fund risk-taking.

The literature on mutual funds find that managers with greater fund ownership are associated with less risk-taking behavior (Ma and Tang, 2014); managerial ownership is also positively related to fund style-adjusted returns (Evans, 2008) and four-factor alphas (Khorana, Servaes, and Wedge, 2007); accounting for fund performance, funds with managerial ownership attract more investor flows compared to those with none (Ma and Tang, 2014). For hedge funds, higher managerial ownership can predict higher future returns (Agar-

[1]For example, under "Ownership" section in Form ADV, a related question is "what is the approximate percentage of the private fund beneficially owned by you and your related persons:"

wal, Daniel, and Naik, 2009); high inside investment is associated with smaller funds, which deliver superior performance as they operate closer to their optimal scale (Gupta and Sachdeva, 2018).

Prior research provides evidence in support of the positive effects of manager's skin-in-the-game on fund risk-taking and performance. However, there is a lack of causal evidence as it is challenging to find an exogenous shock to manager's skin-in-the-game. To complement the earlier studies, I use a unique dataset to establish a causal link between manager's skin-in-the-game and fund risk-taking. The data come from Zulutrade, a major online social trading platform that connects retail traders (fund managers) with investors. I observe public information updated daily that covers trader's profile, historical trades, and individual investor of the fund. In this paper, I investigate the differences of traders with and without skin-in-the-game as well as risks of their managed funds. I use a source of exogenous variation in skin-in-the-game resulting from the platform's exit from the U.S. market to show that removing skin-in-the-game increases managers' risk-taking, thereby providing empirical evidence that skin-in-the-game serves as an important signal of incentive alignment.

I first introduce the design of this study that allows a cleaner inference about the causal relationship between trader's skin-in-the-game and fund risk-taking. On Zulutrade, there are two distinct groups of users: traders and followers (investors). Once a follower follows a trader, the platform copies any new trade made by the trader into follower's brokerage account in real time. The platform charges the follower for each trade copied and then splits the revenue with the trader. One key feature of the platform is that it allows a trader to send out trading signals through a virtual trading account without risking any personal money. However, if the trader chooses to risk her/his own money when trading, Zulutrade puts a prominent dollar ($) badge on the trader's profile. This skin-in-the-game information is public and updated daily by the platform. The trader can risk personal money in one of the two ways: (1) have both a personal follower account and a trader account on the platform and set the follower account to copy trades from the trader account or (2) link her/his own brokerage account with the trader account. The skin-in-the-game trader of Type (1) has a blue dollar badge (hereafter blue trader); the skin-in-the-game trader of Type (2) has a green dollar badge (hereafter green trader). In mid-2016, Zulutrade withdrew from the U.S. market due to regulatory issues and stopped serving U.S. followers[2]. This intervention only affected the traders and followers from the U.S. The U.S. traders who used to have their own follower accounts copy their trader accounts were forced to remove skin-in-the-game as their own follower accounts were closed by the platform. Nevertheless, the U.S. traders can continue sending out trading signals to the remaining non-U.S. followers on the platform after the intervention. This creates an ideal quasi-experiment setting to compare before and after intervention risk-taking of the affected skin-in-the-game U.S. traders with that of the other unaffected skin-in-the-game traders. For measurement precision, I evaluate fund performance and risk-taking in terms of daily net profits. The performance is measured by mean of daily net profits and the risk-taking is measured by kurtosis of daily net profits.

To identify the role of having no skin-in-the-game, I conduct a diff-in-diff analysis. Specifically, I assign the affected U.S. traders into the treatment group. To find a control group, I match each treated U.S. trader to a trader that was un-

[2]Visit `https://www.leaprate.com/2016/10/zulutrade-exits-us-retail-forex-market-withdrawing-nfa-membership-finalizes-30000-settlement/` for more details

affected by the platform's U.S. exit. The matched control traders are blue foreign traders who must always have skin-in-the-game pre- and post- intervention. For the matching, I first estimate propensity score by logistic regression, where I include the following covariates: average fund size, trader's average rank, as well as mean, standard deviation, skewness, kurtosis of daily fund profits over the 8 weeks before the intervention date t (i.e. $t-8$ weeks to $t-1$ week). Then for each of the treated traders, I match one control trader on the estimated propensity score by nearest-neighbor matching. A balance test of covariates reveals that the treated and matched control traders do not differ in any systematic way pre-treatment. Next, I compare how fund performances and risk-taking of the two groups of traders evolve in the subsequent 8 weeks post-intervention. The 8-week time window is chosen as in the sample a follower on average invests in a fund for 8 weeks before leaving. The results suggest that having no skin-in-the-game significantly increases kurtosis of daily profits, which means higher chance of extreme gains or losses. Meanwhile, I find non-significant effect of skin-in-the-game on mean or skewness of daily profits, and weak effect on standard deviation. Overall, having no skin-in-the-game increases trader's tendency to take extreme gains or losses.

In addition to the causal evidence, I conduct a panel regression based design using my full sample from November 2015 to September 2017 to uncover potential drivers of fund risk-taking. There are four separate regressions with mean, standard deviation, skewness, or kurtosis of daily profits as the dependent variable respectively. The main independent variable is trader's skin-in-the-game on day t with other trader fund characteristics as controls. The results suggest that putting skin-in-the-game is associated with lower kurtosis and more positive skewness of daily profits, indicating more conservative trading with less chance of extreme losses. These are in line with the earlier diff-in-diff analysis which instead shows more aggressive trading caused by removing skin-in-the-game. While the full-panel results also indicate worse future performance associated with skin-in-the-game, they are broadly consistent with the diff-in-diff evidence on fund risk-taking. This indicates that the findings from this study could be generalized to a larger population.

This paper has implications for the role of managerial incentives on fund managers' risk-taking behavior. Chevalier and Ellison (1997) have documented a convex flow-performance relationship, in which mutual fund investors reward good performance and do not punish poor performance equally. This in turn can motivate fund managers to strategically shift risk levels to attract additional fund flows, which can lead to worse subsequent abnormal returns (Huang, Sialm, and Zhang, 2011). Ma and Tang (2014) demonstrate that managerial ownership is an important mechanism to reduce mutual fund risk taking. They find that managers with higher personal ownership of the fund or skin-in-the-game engage in less risk-taking, have better Sharpe ratios and attract more flows. I provide evidence that skin-in-the-game reduces managers' risk-taking behavior. In addition, recent studies have highlighted the importance of mandatory disclosure of information regarding fund governance. Khorana, Servaes, and Wedge (2007), Evans (2008), and Cremers et al. (2009) have explored the managerial ownership information provided by the mutual funds, which was newly available after the SEC enacted disclosure rules in the early 2000s. They find higher managerial ownership is associated with better future return. For hedge funds, Brown et al. (2008) and Ozik and Sadka (2015) show that private information about a fund ownership structure may constitute material information for mitigating agency issues. Agarwal, Daniel, and Naik (2009) use managerial ownership to capture managerial incentives in hedge funds, and show higher levels of managerial ownership deliver superior alphas. Recent study by Gupta and Sachdeva (2018) uses Form ADV data to link inside investment to hedge fund returns and shows that insider funds outperform and tend to be smaller, possibly because managers use better strategies in funds with their own private capital and keep operating these funds closer to optimum scale. My paper contributes to this strand of literature by showing that skin-in-the-game as a proxy for fund governance could better inform investors about the fund managers. In my data, fund manager's skin-in-the-game is made public and prominently displayed to all investors and updated daily. I show evidence that changes in skin-in-the-game influence manger's risk-taking behavior, thereby serving as material information for investors.

The findings in this paper are also of interest to regulators. Following the 2008 financial crisis, regulators have been playing a bigger role in protecting investors. For example, to address moral hazard problem endemic to securitization[3], the Dodd-Frank Act's risk retention rule mandates that originators and securitizers retain a 5% interest in their securitizations. This skin-in-the-game rule is intended to align the incentives of securitizers and investors. Regulators could consider similar minimum ownership rules to mutual and hedge fund managers. Moreover, future rules could also aim to increase disclosure requirements for factors relevant to conflict of interest between managers and investors. By shedding light on the positive role of showing skin-in-the-game, I hope that future regulations would focus on more explicit disclosure of information concerning fund management, such as managerial ownership, compensation struc-

[3]Loan originators used to be able to quickly sell loans into securitization pools and not bear any risk on the ultimate performance of the sold loans. This contributed to loosened underwriting standards and riskier loans revealed by the financial crisis.

ture, and conflicts of interest.

The remainder of this paper is organized as follows. Section 1.2 presents the data, variable construction, summary statistics and panel regression results. Section 1.3 provides the main finding on the effect of skin-in-the-game on fund risk-taking. Section 1.4 provides supporting evidence from panel regression of the full sample. Section 1.5 concludes the paper.

1.2 Data

1.2.1 Source

The data are from Zulutrade, one of the major online social trading platforms with over 700,000 registered users focusing entirely on speculative spot forex trading[4]. The users consist of two distinct groups: traders and followers (investors)[5]. The traders make the trading decisions (i.e. long or short a currency pair), and the followers passively copy the exact same trades in real time through Zulutrade platform[6] (see Figure A.1). Any individual can register as a trader without any license requirements. To become a follower, one must have a brokerage account from a broker affiliated with Zulutrade and sign an agreement to authorize Zulutrade's automated trade execution and a Zulutrade markup fee which is incorporated into transaction fees on each trade. In this way, Zulutrade makes money by collecting its share of the trading transaction fees from followers' brokers (1 to 1.5 pips) on each trade (long or short)

[4]See Appendix A.2 for more details about the trading instruments.

[5]An individual can have both a trader and a follower accounts on the platform, but she/he must sign up the two types of accounts separately.

[6]both limit orders and market orders are supported for copying.

copied through Zulutrade. Then Zulutrade splits this revenue to compensate the traders[7].

As a medium between traders and followers, Zulutrade serves two important roles: 1. provide publicly accessible performance details about its traders and 2. provide automated and flexible trader-following settings to followers. For the first role, it has a transparent interface to rank and showcase the top performing individual retail traders. For any trader listed on the platform, the summary performance profile as well as complete historical trade record[8] are open to public view (see Figures A.2, A.3). For the second role, Zulutrade allows a follower to follow a portfolio of traders and customize trade size by trader and by currency pairs. For example, a follower follows Trader A and sets a fixed 0.01 lot size on any EUR/USD trade from Trader A. Then when Trader A opens a long trade on EUR/USD, Zulutrade will automatically open a long EUR/USD trade with 0.01 lot in the follower's brokerage account regardless of Trader A's trade size (see Figures A.4, A.5). The trade is executed in real time by Zulutrade on the follower's behalf.

1.2.2 Sample Construction

The data have been collected from Zulutrade through a scraping program which captures the publicly available trader profile information each day. The data in-

[7]Traders are paid 0.5 pip for each trade (long or short) executed in a follower account. At the end of the month the traders are paid only if they are in profit for the month. They are not paid during a losing month. This rule was introduced by Zulutrade from November 2011, before that traders were paid regardless of a winning or losing month. See Appendix A.3 for more details about the social trading platform business models.

[8]The trading history of a listed trader is available since the time of his/her registration on the platform. A trader cannot selectively modify or hide particular trades since the trading record comes directly from live data feed of trader's brokerage firm untainted by any trader-side record manipulation.

clude the observed information shown in Figures A.2, A.3. The sample period covers from November 2015 to September 2017 with 367 trading days. On average each day I observe about 25,000 listed traders and 13,000 followers with $300 million daily trading volume[9]. Zulutrade provides information at the trader account level. Each trader account has its own profile page associated with a permanent nickname and ID number, but Zulutrade does not disclose the real identity of the trader. So I refer each trader account as one distinct trader, and I treat the fund from all followers following a trader account as one distinct fund[10]. That is each trader is associated with one distinct fund consisting of followers' money. To study fund risk-taking, I subset the sample to include all observations of a trader that manages fund with more than 0 U.S. dollar at any point of time during the sample period. For example, if a trader has a fund with more than $0 at time t, then all observations of the trader before and after time t are included regardless of whether the fund later becomes $0. Over the whole sample period, there are 7,644 unique traders/funds with 1,543,081 trader/fund-day observations in total.

1.2.3 Fund Performance and Risk-taking Measures

I measure fund performance and risk-taking in terms of the daily net profits. Zulutrade provides information about the cumulative fund net profits (after transaction, platform and trader fees) aggregated over all followers in a fund since the fund inception. The daily net profits are then computed by taking first difference of the cumulative profits:

[9] The trading volume only includes followers' trades executed through Zulutrade

[10] One can sign up multiple trader accounts. However, there's no clear way to unravel which accounts belong to the same individual or to group the funds at the real trader level

$Net\ Profits_t = Cumulative\ Profit_t - Cumulative\ Profit_{t-1}$

Though I can also observe the fund size, the fund size and the fund net profits are updated independently with different frequency and there is no public information about the fund net flow. Thus, for measurement precision at the daily level, I evaluate performance in terms of daily profits instead of daily percentage return.

The key fund measures in this paper are the following:

1. Performance measure: mean of daily net profits
2. Risk-taking measure: kurtosis[11] of daily net profits

Additional fund measures include standard deviation and skewness[12] of daily net profits.

1.2.4 Trader's Skin-in-the-game

Zulutrade explicitly shows that a trader has skin-in-the-game by putting a dollar ($) badge next to a trader's profile picture (see Figure A.2). This information is public and updated daily, so I am able to keep track of each trader's change

[11] Kurtosis is defined as excess kurtosis, a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. The normal distribution has a kurtosis of 0. A positive kurtosis means that relative to normal distribution, the data distribution has a sharper peak or more frequent observations close to the mean, and has fatter/heavier tails or more frequent extreme values. A negative kurtosis means that relative to normal distribution, the data distribution has a flatter peak or less frequent observations close to the mean, and has thinner/lighter tails or less frequent extreme values.

[12] Skewness is a measure of the asymmetry of a distribution: a negatively/left skewed distribution has the mass of the distribution concentrated on the right and a longer left tail; a positively/right skewed distribution has the mass of the distribution concentrated on the left and a longer right tail.

in skin-in-the-game over time. This also allows me to compare trader characteristics as well as fund risk and performance between traders^{S} (traders with skin-in-the-game) and traders^{NS} (traders with no skin-in-the-game) each day.

On Zulutrade, a trader can choose whether to put skin-in-the-game whereas all followers have skin-in-the-game, subject to gains or losses in their brokerage accounts when following a trader. To have no skin-in-the-game, a trader links the Zulutrade trader account with a demo (play money) account; when the trader places trades, these trades are entirely simulated for the trader, but the followers follow the same trades with real money.

Within the traders with skin-in-the-game, they can be further differentiated into two groups: green and blue traders. A green trader must link his/her Zulutrade trader account with his/her own real money brokerage account. A blue trader must have both a trader and a follower accounts on Zulutrade; he/she must first link a demo account with Zulutrade trader account and then follow the trader account with his/her own follower account. An advantage of being a green trader is that one can trade without the trading compliance rules[13] placed by the platform on the other types of traders. An advantage of being a blue trader is that a green trader risks personal money at all times, while a blue trader can use his/her own follower account to selectively decide when to risk personal money.

For this study, I group the green and blue traders together into the skin-in-the-game traders as both traders risk real money following their own trading. The amount of skin-in-the-game of each trader is not observed as the platform keeps the balance and equity of trader's brokerage account private. Instead, I

[13] Visit `https://www.zulutrade.com/trader-guide` for details about the trading compliance rules

use a dummy variable $skin$ to indicate whether a trader has skin-in-the-game. Notice that a trader is free to put or remove skin-in-the-game anytime such that a trader$^{\text{NS}}$ today can be a trader$^{\text{S}}$ tomorrow and vice versa, and thus so is the corresponding fund. I observe that 607 unique traders/funds have switched between $^{\text{S}}$ and $^{\text{NS}}$ type in the sample. There are 976 switch events from the 607 traders: 111 switches are from green to $^{\text{NS}}$, 497 from blue to $^{\text{NS}}$, 368 from $^{\text{NS}}$ to blue, and no switch from $^{\text{NS}}$ to green. To account for this time variation of skin-in-the-game for each trader, for each trading day t I re-group the traders/funds into traders/funds$^{\text{S}}$ or traders/funds$^{\text{NS}}$ depending on their skin-in-the-game at t. This avoids the potential issue of assigning a permanent $^{\text{S}}$ or $^{\text{NS}}$ label to each trader/fund.

1.2.5 Summary Statistics

Table 1.1 compares the characteristics of traders$^{\text{S}}$ and traders$^{\text{NS}}$ as well as the performances and risks of their corresponding funds$^{\text{S}}$ and funds$^{\text{NS}}$ head to head. On average 495 traders/funds$^{\text{S}}$ and 1315 traders/funds$^{\text{NS}}$ are reported daily in the data. I follow the Fama and MacBeth (1973) approach to first calculate the cross-sectional mean of each variable for traders$^{\text{S}}$ and traders$^{\text{NS}}$ respectively each day, and then report the respective time series mean of the daily means of each variable in "Mean$^{\text{S}}$" and "Mean$^{\text{NS}}$" columns of Table 1.1.

In Panel A of Table 1.1, notice that on average Trader$^{\text{S}}$ appears to be less popular than Trader$^{\text{NS}}$. Trader$^{\text{S}}$ manages a smaller fund than Trader$^{\text{NS}}$ - \$18,715 versus \$19,334, though the difference is not statistically or economically significant. Trader$^{\text{S}}$ also has fewer followers, worse followers' ratings, and worse rank. This might be explained by Trader$^{\text{S}}$'s shorter account age on the platform (76 weeks vs Trader$^{\mathrm{NS}}$'s 96 weeks). However Trader$^{\mathrm{S}}$ seems more conservative in trading, holding trades in shorter duration (21% shorter than Trader$^{\mathrm{NS}}$' average trade duration) and keeping fewer trades open at the same time. In addition, Trader$^{\mathrm{S}}$ does 16% better for the worst historical trade and has 56% lower historical maximum drawdown of open trading positions in terms of pips[14], implying that Trader$^{\mathrm{S}}$ generally fares better than Trader$^{\mathrm{NS}}$ in a worst-case scenario.

In Panel B of Table 1.1, I focus on the fund performance and risk-taking measured in terms of daily profits. I first get the daily profits at day t respectively for the funds which were of $^{\mathrm{S}}$ or $^{\mathrm{NS}}$ type at $t-1$. I then report the time series means of different daily cross-sectional statistics of the daily profits for Funds$^{\mathrm{S}}$ and Funds$^{\mathrm{NS}}$ respectively. Notice that Funds$^{\mathrm{S}}$ and Funds$^{\mathrm{NS}}$ have mean daily losses of \$19.19 and \$17.90 with standard deviations of \$854.27 and \$821.98 respectively without any significant differences in means. They also have close to 0 median profits. These are consistent with the view that forex trading has zero expected return due to the zero-sum game nature[15]. Since the performances are measured after fees, the zero return view still applies despite the mean daily losses. To compare the distributions of daily profits between Funds$^{\mathrm{S}}$ and Funds$^{\mathrm{NS}}$ in depth, I also use Max, Min, $Skewness$ and $Kurtosis$. Both Funds$^{\mathrm{S}}$ and Funds$^{\mathrm{NS}}$ have negatively skewed profits meaning that they have mostly consistent small profits but with occasional large losses. However, Funds$^{\mathrm{S}}$' profits are significantly less negatively skewed which means less chance of extremely losses. Also Funds$^{\mathrm{S}}$' kurtosis is 50% smaller than that of Funds$^{\mathrm{NS}}$. This means Funds$^{\mathrm{S}}$ exhibit significantly less frequent extreme gains or losses. For the extreme cases,

[14]A pip (price interest point) is the smallest price movement in exchange rate. It is used to calculate the value change of a forex trading position. See Appendix A.4 for an example.

[15]In forex trading, every profitable position is matched by an opposite losing position. For example, if one bets euro will be stronger than dollar, there must be another who bets dollar will be stronger than euro in order for the trade to occur. In this respect, forex trading does not generate a net profit to reward all holders of currency risk.

Max and Min statistics show that Funds^{S} have 24% smaller gains on average in the best cases, but lose 32% less on average in the worst cases than Funds^{NS}. Moreover, for both fund types the worst case losses are more than double the best case gains, and extreme gains and losses have much bigger magnitudes than mean profits. This calls for attention about the kurtosis of profits which deals with tail risk or risk of rare events. In particular, fat tails, the negative or left tail to be exact, entail higher chance of extreme losses, which could ruin investors' wealth when realized.

1.3 Evidence from a Diff-in-Diff Analysis

The summary statistics show that having skin-the-game can decrease the tail risk, making funds less prone to extreme gains or losses. However, this finding could be influenced by potential endogeneity problems. One concern is that a trader has private information about fund performance that influences putting or removing skin-in-the-game. For example, a trader may start with no skin-in-the-game and strategically choose to put skin-in-the-game only after a sequence of big wins or losses. Assume the trader stays with the original trading strategy. If the performance exhibits reversion to the mean, this can lead to downward bias of the effect of skin-in-the-game on the tail risk. Another concern is there might exist some unobserved systematic differences between $\text{traders}^{\text{S}}$ and $\text{traders}^{\text{NS}}$.

To get a cleaner inference about the causal effect of skin-in-the-game on performance and risk-taking, I conduct an event study exploiting the platform's exit from the U.S. market. During late May to early July in 2016, Zulutrade withdrew from the U.S. market and terminated services to all U.S. follower accounts[16]. However, any existing U.S. trader could continue to trade on the platform and collect performance fees from the remaining non-U.S. followers. As defined in Section 1.2.4, a blue trader is one type of traders with skin-in-the-game who must have both a trader account and a follower account on the platform. Now that no U.S. followers are allowed, all U.S. traders that used to be blue no longer have skin-in-the-game and their blue dollar badges were removed accordingly as their own follower accounts were closed by the platform. However, the traders that were using a U.S. broker before the intervention could still use the broker with the same account as before, but they could not make new changes with the U.S. broker (i.e. change from a demo account to a live account vice versa) as the platform ended affiliation with the U.S. brokers as part of the U.S. exit. So the U.S. green traders were still green and had skin-in-the-game despite the intervention. However, former blue U.S. traders could not switch to be green since they would need a real brokerage account of a U.S. broker to replace their current demo accounts, which was not possible after the intervention. In other words, the former U.S. blue traders had no choice to put skin-in-the-game again and none of them were observed to be green afterwards.

This platform intervention provides an ideal quasi-experiment setting where the exogenous variation is that the U.S. blue traders were forced to remove skin-in-the-game since their own follower accounts were terminated by the platform. This allows me to identify the role of having no skin-in-the-game by comparing pre- and post-U.S. exit trader/fund characteristics. There were 22 traders/funds affected by the intervention. I put them into the treatment group. In the next section, I use matching to construct a control group. The idea is to match on

[16]See footnote 2

some key confounding factors to make the treatment and control groups closely comparable so that the only systematic difference between the treatment and control groups will be the treatment itself.

1.3.1 Matching Process

For each of the 22 treated traders, I match one control trader using nearest-neighbor matching. The distance is measured by a propensity score with logistic regression. The control traders are blue foreign traders as they were unaffected by the platform's U.S. exit, and they must always have skin-in-the-game pre- and post-intervention. To ensure the treated and control traders have similar fund performances, risk-taking, and platform characteristics pre-intervention, I include mean, standard deviation, skewness, kurtosis of daily fund profits as well as average fund size and trader's average rank over the 8 weeks before the intervention date t (i.e. t-8 weeks to t-1 week) in propensity score estimation. I choose to evaluate the performance and risk-taking over 8 weeks (roughly two calendar months) so that there are more trading activities captured with more variability in the daily profits for the treated funds. Overall 16 treated traders were matched, and the rest were unmatched as they were inactive 8 weeks before the intervention. As shown in Panel A of Table 1.2, the covariates are fairly balanced across the treated and matched control traders.

1.3.2 Results

I show the diff-in-diff regression results in Panel B of Table 1.2. I examine fund performance and risk-taking in four separate regressions with mean, standard deviation, skewness, or kurtosis of daily profits as the dependent variable respectively. I use the regression framework as follows:

$$y_{i,t} = \alpha_i + \alpha_t + \beta_1(treatment_i \times post_{i,t}) + \beta_2 fundsize_{i,t} + \beta_3 rank_{i,t} + \epsilon_{i,t} \quad (1.1)$$

where $y_{i,t}$ is the dependent variable, which is mean, standard deviation, skewness or kurtosis of daily profits of fund i (measured over an eight-week rolling window); α_i and α_t capture fund and calendar day fixed effects, respectively; $treatment_i$ is treatment dummy variable (1: fund i is a former U.S. blue fund; 0: fund i is a control fund); $post_{i,t}$ is time dummy variable (1: date t is post-intervention; 0: date t is pre-intervention); $fundsize_{i,t}$ and $rank_{i,t}$ account for average fund size and average trader rank over the previous 8 weeks, respectively. The main coefficient of interest is β_1. The sample includes fund-day observations of 16 treated funds and 16 control funds from March 1 to September 1 2016.

The results show that having no skin-in-the-game significantly increases kurtosis of daily profits, whereas there is no significant effect on mean or skewness and weak effect on standard deviation of daily profits. As shown in Figure 1.1, treated and matched control traders have parallel trends in their kurtosis of daily fund profits before the intervention but the treated traders exhibit a big spike in kurtosis after the intervention. Therefore, having no skin-in-the-game makes extreme gains/losses more likely even though it may not significantly affect performance or other measures. This makes traders' skin-in-the-game potentially important in accessing the tail risk of funds.

1.4 Full Sample Evidence

To check how well the earlier results generalize, I proceed with a panel regression design with a bigger sample. I use my full sample from November 2015 to September 2017 with the following regression framework:

$$y_{i,t+n} = \beta skin_{i,t} + \Gamma \cdot X_{i,t} + \alpha_i + \eta_t + \epsilon_{i,t} \tag{1.2}$$

where $y_{i,t+n}$ is the dependent variable, which is mean, standard deviation, skewness or kurtosis of daily profits of fund i from $t+1$ trading day to $t+n$ weeks (a rolling window of n weeks); the independent variable of interest $skin$ dummy equals 1 if the trader of fund i has skin-in-the-game on day t and 0 otherwise; $X_{i,t}$ is a set of controls including fund size and profits, trader's rank, rating, trade duration, account age, win ratio, and pips on day t. To control for the period when a fund is potentially defunct, I also include a $no\text{-}asset$ dummy which equals one if a trader has zero asset under management on day t and zero otherwise. To capture individual and time fixed effects, I include α_i for trader or fund fixed effects, and η_t for day fixed effects. For testing significance, I use the robust t-statistics clustered by day.

I present the regression results for a rolling window of four weeks in Table 1.3. The results suggest putting skin-in-the-game lowers kurtosis and increases positive skewness of daily profits. This resembles a more conservative trading style with less frequent extreme losses. Surprisingly for performance, putting skin-in-the-game appears to decrease future mean daily profits. Overall, the results show that skin-in-the-game makes trading more conservative even though it may come at the expense of worse mean daily profits. The results are also robust to different rolling windows. Therefore, the full-sample results are broadly consistent with the earlier diff-in-diff findings which could be generalized to a larger population.

1.5 Conclusion

As mutual funds and hedge funds continue to grow, it is increasingly urgent to determine the factors that influence fund manager trading behavior. This paper studies the relation between a manager's skin-in-the-game and the risk-taking of the associated fund using a unique dataset from a social trading platform that has time variation of skin-in-the-game at a daily frequency. I explore the differences of traders with and without skin-in-the-game as well as the risk-taking of their managed funds. My paper contributes to the literature by providing empirical evidence of the causal effect of skin-in-the-game on fund risk-taking.

In particular, I use a source of exogenous variation in manager's skin-in-the-game for identification. This comes from an intervention event in which the platform exited from the U.S. market forcing some U.S. traders to have no skin-in-the-game. I find that skin-in-the-game has significant impact on trader's incentive for risk taking. The results show strong evidence that having no skin-in-the-game motivates traders to take extreme gains or losses more frequently as revealed by the higher kurtosis of daily profits. Thus, not putting any skin-in-the-game encourages risk taking as the traders could aim for extreme gains without sharing the downside risk with investors.

Overall, my findings suggest that putting no skin-in-the-game can motivate excessive risk-taking. Therefore, requiring skin-in-the-game is important for mitigating agency conflicts in managed funds. I believe these results have broad regulatory implications regarding mandatory minimum skin-in-the-game and disclosure policies for mutual funds and hedge funds.

1.6 Figures

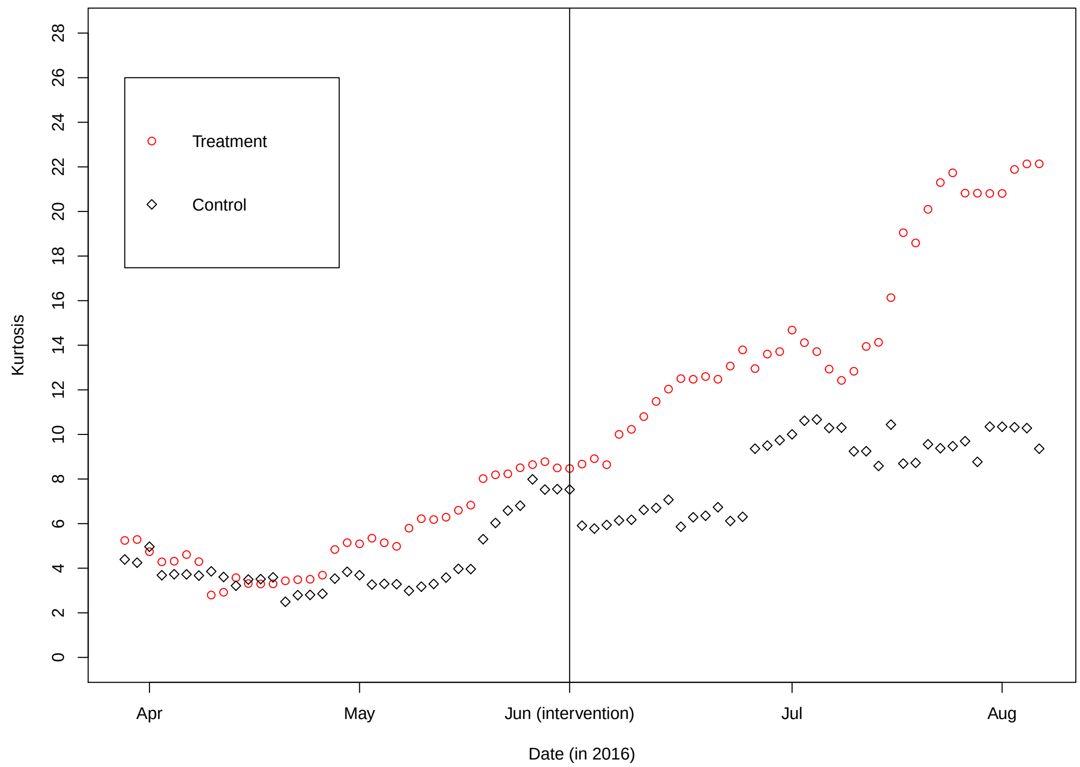

Figure 1.1: Funds' kurtosises around the platform intervention

This figure plots the kurtosis of daily profits over time for the treated funds and the matched control funds over the sample period. The intervention in June 2016 forced all U.S. blue traders to remove skin-in-the-game (treatment) due to the platform's exit from the U.S. market. Kurtosis at date t is calculated over the fund daily profits using an eight-week rolling window.

1.7 Tables

Table 1.1: Summary Statistics for Traders/Funds

This table reports the summary statistics for the sample covering from November 2015 to Sepetember 2017 with $N = 367$ trading days. I first calculate the cross-sectional mean of each variable for traders with skin-in-the-game (denoted by the superscript ^{S}) and traders with no skin-in-the-game (denoted by the superscript ^{NS}) respectively on each day and then I report the time series mean of the daily means for the two types of traders respectively in "Mean^{S}" and "Mean^{NS}" columns of this table; in "Diff" column, I report the mean of time series differences for each variable between the two types of traders; I use the paired t-procedure to test the significance of the mean of the differences and adjust the t-statistic following Newey and West (1987) method at six lags to overcome serial correlation and heteroskedasticity; the corresponding P-value is reported in "P-value" column with ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively. Variable definitions: $fundsize$ is the aggregate balance (in U.S. dollars) of all followers following a trader. $follower$ is the number of followers. $rating$ is the average trader ratings out of 5 rated by the followers. $rank$ is Zulurank (the smaller the better) on the platform. $trades$ is the total number of historical trades. age is the length of trader's trading history (in weeks) on the platform. $tradedur$ is the average duration (in days) of the historical trades. win is the win percentage of the historical trades. $maxopentrades$ is the max number of trades that have ever been kept open at the same time. $pips$ is the total pips made from historical trades (see Appendix A.4 for more details). $pipsMDD$ is the historical maximum drawdown of open trading positions in pips. $avgpips$, $worst$-$trade$, $best$-$trade$ are the average, worst, and best of all historical trades in terms of pips respectively.

	Mean^{S}	Mean^{NS}	Diff	P-value
Panel A: Daily trader profile information				
fundsize ($)	18715.46	19333.88	−618.42	0.353
followers	24.68	26.96	−2.28	0.001***
rating	2.08	2.20	−0.12	0.000***
rank	13437.04	9845.07	3591.97	0.000***
trades	968.53	1292.44	−323.91	0.000***
age (weeks)	76.40	95.20	−18.80	0.000***
tradedur (days)	3.82	4.83	−1.01	0.000***
win (%)	69.34	76.17	−6.82	0.000***
maxopentrades	20.22	26.24	−6.02	0.000***
pips	1140.62	20078.08	−18937.46	0.000***
pipsMDD	11396.85	25877.71	−14480.86	0.000***
avgpips	3.42	28.81	−25.40	0.000***
worst-trade (pips)	−1241.41	−1474.15	232.74	0.000***
best-trade (pips)	589.95	1052.86	−462.92	0.000***
Panel B: Daily fund profit ($)				
Max	6281.68	8281.97	−2000.29	0.039**
Min	−13322.47	−19696.89	6374.41	0.023**
Mean	−19.19	−17.90	−1.28	0.833
Median	0.01	−0.03	0.04	0.731
Percentile-25	−1.80	−12.10	10.29	0.000***
Percentile-75	2.02	10.94	−8.92	0.000***
SD	854.27	821.98	32.29	0.748
Skewness	−3.32	−6.47	3.15	0.002***
Kurtosis	204.72	412.04	−207.32	0.000***

Table 1.2: Fund Performance and Risk-taking around the Platform Intervention

Panel A: Pre-treatment statistics: this panel reports balance test of covariates after the match. The treatment group consists of all the former U.S. blue traders who were forced to remove skin-in-the-game by the platform intervention. The control group consists of the foreign blue traders who had skin-in-the-game pre- and post-intervention. I match a treated trader to a control trader using one-to-one nearest-neighbor matching. The distance is measured by a propensity score with logistic regression. I include mean, standard deviation, skewness, kurtosis of daily fund profits as well as average fund size and trader's average rank over the 8 weeks before the intervention date t (i.e. t-8 weeks to t-1 week) in propensity score estimation.

	Treatment	Control	Diff	t-stat
Mean	−231.77	−83.94	−147.83	−0.64
SD	1,433.25	363.66	1,069.59	0.75
Skewness	−0.43	−0.15	−0.28	−0.26
Kurtosis	9.85	10.35	−0.51	−0.15
fundsize	33,816.01	10,297.33	23,518.68	0.72
rank	15,828.85	13,161.86	2,666.99	0.87

Panel B: Regression results: this panel reports the result of diff-in-diff regression of fund performance and risk-taking measures around the platform intervention. The sample includes fund-day observations of 16 treated funds and 16 control funds from March 1 to September 1 2016. In (1) to (4) columns, I use mean, standard deviation, skewness, or kurtosis of daily profits measured over an eight-week rolling window as the dependent variable respectively. The performance is measured by mean of daily profits and the risk-taking is measured by kurtosis of daily profits. $treatment$ is a dummy variable which equals one for a former blue U.S. trader and zero otherwise; $post$ is a time dummy variable which equals one if date t is post-intervention and zero otherwise. The main explanatory variable $treatment * post$ is an interaction term between the two dummy variables. All regressions include trader and calendar day fixed effects with the robust t-statistics clustered by day reported in parentheses.

	Dependent variable:			
	Mean	SD	Skewness	Kurtosis
	(1)	(2)	(3)	(4)
treatment*post	2.07	-133.84^{*}	0.03	2.71^{***}
	(19.85)	(69.93)	(0.13)	(0.75)
fundsize	-0.01^{***}	0.02^{***}	-0.00^{***}	0.00^{**}
	(0.00)	(0.00)	(0.00)	(0.00)
rank	0.01^{***}	-0.03^{***}	0.00	0.00^{*}
	(0.00)	(0.00)	(0.00)	(0.00)
fund FE	Y	Y	Y	Y
day FE	Y	Y	Y	Y
Observations	2,984	2,984	2,748	2,748
R^2	0.55	0.66	0.27	0.56
Adjusted R^2	0.52	0.64	0.23	0.54

Table 1.3: Fund Performance, Risk-taking and Trader's Skin-in-the-game

This table reports panel regression result of trader's skin-in-the-game on fund performance and risk-taking. The sample consists of trader fund-day observations from November 2015 to September 2017 covering 367 trading days. There are four separate regressions with mean, standard deviation, skewness, or kurtosis of daily profits as the dependent variable respectively. The fund performance is measured by mean of daily profits and the risk-taking is measured by kurtosis of daily profits. For a fund on trading day t, each dependent variable in this panel is computed using the future daily profits of the fund starting from $t+1$ trading day up to $t+4$ weeks (i.e. a four-week rolling window). The main independent variable $skin$ dummy equals one if a trader has skin-in-the-game on day t and zero otherwise. $profit$ is fund's daily profit on day t. *no-asset* dummy equals one if a trader has zero asset under management on day t and zero otherwise. The other control variables are defined in Table 1.1. All regressions include trader and calendar day fixed effects with the robust t-statistics clustered by day reported in parentheses.

	Dependent variable:			
	Mean	SD	Skewness	Kurtosis
	(1)	(2)	(3)	(4)
skin	-17.23^{***}	50.38^{***}	0.04^{**}	-0.06^{**}
	(2.62)	(6.87)	(0.02)	(0.03)
fundsize	-0.00^{***}	0.01^{***}	-0.00^{***}	0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
profit	0.00	0.01	0.00^{**}	-0.00
	(0.00)	(0.01)	(0.00)	(0.00)
rank	0.00^{**}	-0.00^{***}	0.00^{***}	-0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
rating	-4.70^{***}	18.34^{***}	-0.14^{***}	-0.13^{***}
	(1.14)	(2.81)	(0.01)	(0.01)
tradedur	-0.01^{*}	0.02	0.00	-0.00^{***}
	(0.00)	(0.02)	(0.00)	(0.00)
weeks	-0.02	-0.27^{***}	0.00^{**}	-0.00^{***}
	(0.01)	(0.05)	(0.00)	(0.00)
win	-0.61^{***}	1.89^{***}	-0.01^{***}	0.05^{***}
	(0.12)	(0.40)	(0.00)	(0.01)
pips	-0.00^{***}	0.00^{***}	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
no-asset	-5.03^{***}	46.12^{***}	0.28^{***}	-1.19^{***}
	(0.84)	(3.92)	(0.01)	(0.03)
trader FE	Y	Y	Y	Y
day FE	Y	Y	Y	Y
Observations	1,434,172	1,426,481	894,496	894,496
R^2	0.15	0.30	0.20	0.45
Adjusted R^2	0.15	0.30	0.19	0.44

CHAPTER 2

EXCESSIVE EXTRAPOLATION OF INDIVIDUAL FUND INVESTORS

2.1 Introduction

A growing body of literature suggests that investors extrapolate past stock returns: investors' expectations about a stock's future return are positively correlated with the stock's recent past returns[1]. While the literature has devoted primary attention to examining extrapolative beliefs with respect to stock market, there is little large-sample evidence about the return beliefs of individual investors to delegated assets i.e. investment funds, or about how investors form extrapolative beliefs besides just from raw past performance. In this paper, I provide novel micro-level evidence that accounting for past performance and risk, individual investors also rely on consistency of past performance in making fund withdrawal decision. That is investors will stay with (leave) the fund if recent performance is consistently positive (negative), even though such extrapolative beliefs do not help the investors' future performance. Furthermore, investors from more developed countries are more subject to extrapolation.

I begin by introducing the data that allow me to examine the existence, determinants and performance consequences of individual fund investors' extrapolative beliefs. The data come from Zulutrade, a major online social trading platform that connects retail traders (fund managers) with retail followers (investors) for trading spot forex. Once an investor follows a trader, the platform copies and executes trades made by the trader into investor's brokerage account in real time. The platform is partnered with investor's brokerage firm, and gen-

[1]For example, Benartzi (2001); Greenwood and Shleifer (2014); Da, Huang, and Jin (2018)

erates revenue by receiving fees from the brokerage firm for each trade executed on investor's behalf, then the platform splits the revenue with the trader. It is simple to draw connections between Zulutrade setting and conventional investment funds: each Zulutrade trader is associated with a distinct fund consisting of investors' money; similar to an actively managed mutual fund, Zulutrade trader as the fund manager decides which trades to make on investors' behalf.

Zulutrade is very transparent about each investor's performance, which allows me to track fund investment, withdrawal and performance at individual investor level. In particular, I observe all investors of a fund from trader's public profile, then I gather all historical trades from each investor of the fund. The historical trades are from investor's public profile, where it shows each trade transaction timestamped and linked with the identity of trader. These trade transactions come from investor's brokerage account. I construct the main sample based on the investors' historical trading records, where for each investor-fund pair I aggregate the trades over the investment period into investor-fund-day observations. This allows me to track daily performance of fund investment at individual investor level. The constructed sample spans 9 years from November 2007 to December 2016. Overall, there are about 9.6 million investor-fund-day observations involving about 21 thousand funds and 17 thousand investors representing more than 160 countries. To my knowledge, this is the only setting which provides a large sample of individual fund investors at a global scale to analyze extrapolative beliefs.

I next examine the existence and determinants of extrapolative beliefs. Specifically, I test the hypothesis that investor expects high (low) performance if past performance is consistently positive (negative), which predicts that an investor will stay with (withdraw from) the fund if recent performance is consistently positive (negative). I model investor's withdrawal decision by survival analysis. The model includes multiple observations per investor-fund pair, one for every trading day until the investor's withdrawal from the fund. The dependent variable is binary indicating whether investor i withdraws from fund j on date t. The main independent variables are two separate daily measures of past performance consistency: the consistency of gains and consistency of losses of fund investment over the past two weeks. In addition to finding that investors extrapolate past performance, I find that investors are more (less) likely to withdraw when experiencing recent consistent losses (gains) after controlling for past performance and risk. This evidence is consistent with extrapolative beliefs, in that investor's expectation about the fund future performance is strongly positively related to its recent past performance and its performance consistency. The findings also highlight that extrapolation depends significantly on past performance consistency in addition to raw performance and risk.

Then I examine whether extrapolation is related to investor's future performance. I use OLS panel regression to test for the relation between past performance and future performance of fund investment. I find that neither consistency of gains nor consistency of losses significantly predicts future performance of fund investment, while high recent raw performance significantly predicts low future performance. Overall, the findings suggest that investors extrapolate based on past performance consistency in addition to past performance, even though past performance and its consistency do not positively predict the future performance. This provides evidence that investors over-extrapolate, in that the average belief of investors is not positively related to the subsequent realized return, suggesting that the extrapolative beliefs are incorrect.

Finally, I examine extrapolation and performance predictability within subsamples of OECD and non-OECD investors respectively[2]. I find that fund withdrawals of the investors from OECD member countries are more influenced by consistency of either good or poor performance as well as by raw past performance than those of the other investors. This suggests that investors from more developed countries are more subject to extrapolation. Nonetheless, the results from performance predictions suggest that both groups of investors do not benefit from extrapolation.

My paper contributes to the literature on investor extrapolation, which posits that investors' expectation of the future return of an asset is a positive function of the asset's recent past returns. Such assumption that investors have extrapolative beliefs is one of the most recognized ideas in behavioral finance. Theoretical papers have presented models which show that return extrapolation helps explain a wide range of facts about asset prices including excess volatility, momentum, and bubbles (Barberis et al., 2015, 2018; Jin and Sui, 2018). Empirical papers have primarily tested extrapolation models by using survey expectations of investors about future stock market returns, and provide evidence in support of extrapolation: for the aggregate stock market, Greenwood and Shleifer (2014), Cassella and Gulen (2018); at individual stock level, Da, Huang, and Jin (2018) who use data from Forcerank, a crowdsourcing platform for ranking stocks. Moreover, Ertan et al. (2017) use brokerage data on individual investors to show that individual investors extrapolate stock returns around earnings announcements. In addition, Andonov and Rauh (2018) document that ex-

[2]The Organisation for Economic Co-operation and Development (OECD) is a group of 36 member countries that promote economic and social policies. The members of the OECD are mostly developed countries with high-income economies

trapolative beliefs also affect target asset allocations for institutional investors. While the literature has established extrapolation in the case of investors to non-delegated assets like stocks, my study is the first that I am aware of to make this determination for investors of delegated assets i.e. fund investors who are geographically diverse.

My findings make two main contributions. First, I show that investors over-extrapolate fund performance. Their fund withdrawal decisions reflect extrapolative beliefs that fund performance trends will persist, but non-positive effect on subsequent performance demonstrate that these beliefs are incorrect. Second, I show that extrapolation is not due exclusively to raw performance since investors also consider the consistency of performance when they extrapolate. Both contributions add to our understanding of prior evidence supporting investor extrapolation. Prior literature has established that stock market participants tend to over-extrapolate (Greenwood and Shleifer, 2014; Ertan et al., 2017; Da, Huang, and Jin, 2018), and I offer evidence that over-extrapolation is also evident for fund investors. Sirri and Tufano (1998) show that fund investors infer manager skill from past returns and therefore chase returns; in the model of Berk and Green (2004), mutual funds face decreasing returns to scale so that when more funds flow in, past performance does not predict future performance. My findings support the interpretation that trend chasing in mutual funds is not a rational strategy, and demonstrate this at individual fund investor level. Moreover, the literature on investor experience suggests that individual experiences of macroeconomic shocks affect investors' expectations Malmendier and Nagel (2011). I demonstrate that investors' personal experiences of investment performance consistency is also an important factor in forming expectations.

The remainder of this paper is organized as follows. Section 2.2 presents the data, variable definitions, sample construction, and summary statistics. Section 2.3 explains the performance consistency measure and survival analysis of investor's withdrawal. Section 2.4 studies the existence and determinants of extrapolative beliefs, examines the relation between extrapolation and investor's future performance, as well as compares the results for OECD and non-OECD investors. Section 2.5 concludes the paper.

2.2 Data

2.2.1 Sources

The primary data source is Zulutrade, one of the major online social trading platforms with over 700,000 registered users focusing mainly on speculative spot forex trading[3]. The users consist of two distinct groups: traders and followers (investors)[4]. The traders make the trading decisions (i.e. long or short a currency pair), and the followers passively copy the exact same trades in real time through Zulutrade platform[5] (see Figure A.1). Any individual can register as a trader without any license requirements. To become a follower, one must have a brokerage account from a broker affiliated with Zulutrade and sign an agreement to authorize Zulutrade's automated trade execution and a Zulutrade markup fee which is incorporated into transaction fees on each trade. In this way, Zulutrade makes money by collecting its share of the trading trans-

[3]See Appendix A.2 for more details about the trading instruments.

[4]An individual can have both a trader and a follower accounts on the platform, but she/he must sign up the two types of accounts separately.

[5]both limit orders and market orders are supported for copying.

action fees from followers' brokers (1 to 1.5 pips) on each trade (long or short) copied through Zulutrade. Then Zulutrade splits this revenue to compensate the traders[6].

As a medium between traders and followers, Zulutrade serves two important roles: 1. provide publicly accessible performance details about its traders and 2. provide automated and flexible trader-following settings to followers. For the first role, it has a transparent interface to rank and showcase the top performing individual retail traders. For any trader listed on the platform, the summary performance profile as well as complete historical trade record[7] are open to public view (see Figures A.2, A.3). For the second role, Zulutrade allows a follower to follow a portfolio of traders and customize trade size by trader and by currency pairs. For example, a follower follows Trader A and sets a fixed 0.01 lot size on any EUR/USD trade from Trader A. Then when Trader A opens a long trade on EUR/USD, Zulutrade will automatically open a long EUR/USD trade with 0.01 lot in the follower's brokerage account regardless of Trader A's trade size (see Figures A.4, A.5). The trade is executed in real time by Zulutrade on the follower's behalf.

The Zulutrade follower data are collected through a scraping program which captures the publicly available follower profiles. On Zulutrade, a follower profile is created once the follower links a brokerage account to Zulutrade. The profile has a permanent ID and stays updated with the corresponding broker-

[6]Traders are paid 0.5 pip for each trade (long or short) executed in a follower account. At the end of the month the traders are paid only if they are in profit for the month. They are not paid during a losing month. This rule was introduced by Zulutrade from November 2011, before that traders were paid regardless of a winning or losing month. See Appendix A.3 for more details about the social trading platform business models.

[7]The trading history of a listed trader is available since the time of his/her registration on the platform. A trader cannot selectively modify or hide particular trades since the trading record comes directly from live data feed of trader's brokerage firm untainted by any trader-side record manipulation.

age account with just a few minutes of delay. In the profile, there is a list of entire historical trades since the follower linked the brokerage account to the platform. Each trade is timestamped and linked with the identity of whom opened the trade (i.e. by the follower or by the follower's chosen traders; see Figure A.5). By default, the profile is public which allows the follower to share trading records with peers. To gather the follower profiles, the scraping program first visits the profile page of each publicly listed trader; then it compiles the follower list under the followers section in each trader's profile, where Zulutrade lists all corresponding followers (including those with private profiles) of the trader; lastly, it visits each publicly available follower's profile to obtain follower's entire historical trading records. For follower demographics, Zulutrade reveals only follower's country and brokerage firm publicly, which are also included in the scraped data.

Data on daily foreign exchange rates are obtained from Thomson Reuters. I use the rates at 5pm New York time every trading day.

2.2.2 Investor, Fund Investment and Withdrawal

To put it in financial terms, each Zulutrade trader is associated with one distinct fund consisting of money from the followers[8]. The U.S. dollar balances aggregated across all followers of the trader are the total asset under management (AUM). Though the platform shows AUM publicly for each fund and lists all corresponding followers, it does not show any follower's balance amount and it does not show if a follower is actively following trader's trades.

[8]One can sign up multiple trader accounts. However, there is no clear way to unravel which accounts belong to the same individual or to group the funds at the real trader level as Zulutrade does not disclose the real identity of the trader.

To address these issues and ensure measurement precision, I determine if a follower invests in a fund at a particular time using follower's historical trading records. I judge investment and withdrawal events by the followed trades from the trader. A follower is said to become an investor of a fund at the opening time of his/her first followed trade from the trader of the fund. The follower stays as an investor if he/she is actively following the trader's trades. I base follower's activeness in following the trader on the number of followed trades each day, and I treat no followed trades from the trader over 21 consecutive trading days (one calendar month) or more as an indicator for follower's withdrawal from the fund. Specifically, a follower is said to withdraw from a fund each time if there is at least a 21 consecutive trading day gap, in which the follower neither has any pre-existing trade nor copies any new trade from the trader. The follower is said to re-invest in the fund if there is any new followed trade after the prior withdrawal.

2.2.3 Sample Construction

The main sample is based on the followers' historical trading records, which were scraped daily in the period of November 2015 to December 2016. As the main trading instruments are spot forex and precious metals on Zulutrade, followers that have only traded spot currencies, gold or/and silver are included in the analysis[9]. Overall, there are 19.2 million trades, 17 million of which were opened by traders with the rest opened by followers themselves. As this paper focuses on investor's behavior to performance of fund investment, I direct my

[9]There are 72 spot currency pairs and 3 spot precious metal pairs included. The excluded trading instruments include oil, natural gas, equity indices, and cryptocurrencies. About 7% of the followers are excluded.

attention to the 17 million trades that were opened by traders.

To construct the main sample, I use the following activeness approach defined previously to determine the period in which a follower is an investor of a fund, and then aggregate the trades over the period into investor-fund-day observations. I aim to evaluate daily investment performance in each fund at individual investor level. As Zulutrade does not show follower's account balance publicly, I measure performance in terms of daily net profits (after transaction, platform and trader fees) instead of daily percentage return. To this end, I first group trades on each trading day by investor fund pair. Then summing over all the trades within an investor fund pair, I obtain the total unrealized net profits, where I use prices from Thomson Reuters to adjust for mark-to-market profits on each day's pre-existing trades. The constructed sample spans about 9 years with the earliest investor fund pair appears in November 2007 with the latest one up to end of December 2016. In total, there are 9,622,281 investor-fund-day observations which include 291,724 unique investor-fund pairs consisting of 17,033 investors and 21,133 funds.

2.2.4 Summary Statistics

An investor in the data holds a fund for 44 trading days (about 2 calender months) on average. The sample used in this paper is restricted to include investor-fund pairs that last at least 10 trading days (2 calendar weeks) because investors need sufficient time with a fund to form future performance expectation, and very quick withdrawals might be driven by non-performance, non-fund related factors.

Table 2.1 provides some summary statistics about the investor-fund pairs in the sample. On average, during investor's fund holding period, the performance consistency of gains ($PosPC$) is 0.12 while the performance consistency of losses ($NegPC$) is 0.05 measured over a rolling window of past 10 trading days. Variables related to reverse disposition effect are $gain\_all$ and $gain\_10days$, which indicate if investor's cumulative net profits from fund are positive since the first investing date and the past 10 trading days, respectively. The value of $gain\_all$ means that the average fraction of investor-fund-day observations at a gain is 0.53 using the first investing date of the fund as the reference period. The performance and risk of fund investment are measured by the mean, standard deviation, and skewness of investor's daily net profits from fund over the past 10 trading days. On average, the daily mean profits from a fund are -$1.51, with $47 standard deviation and positive skewness. Variables related to investor's exposure to a fund are $trades$ and $position\_size$, which are investor's average number of daily open trades and average trading position size from investing in fund measured over the past 10 trading days, respectively. Since fund withdrawal decision is potentially related to whether a follower has invested in a fund before, the table also includes the number of times of re-investment ($reinvest$), if its value is greater than 0, it means the follower has invested in a fund more than once.

2.3 Methodology

2.3.1 Performance Consistency Measure

To quantify the performance consistency (PC) of a fund investment, I follow a methodology motivated by Da, Gurun, and Warachka (2014). In particular, I use signed versions of PC to examine the consistency of gains and consistency of losses separately, denoted by $PosPC$ and $NegPC$. These two measures are defined below using daily unrealized profits made by the fund for the investor:

$$PosPC = \begin{cases} \%pos - \%neg & \text{if } cumprofit > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

and

$$NegPC = \begin{cases} \%neg - \%pos & \text{if } cumprofit < 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

where $cumprofit$ denotes the cumulative unrealized profits during the formation period; %pos and %neg denote the percentage of days during the formation period with positive and negative unrealized profits, respectively. $PosPC$ or $NegPC$ takes on a value between 0 and 1; the larger the value is, the more consistent are past gains or losses. Notice that $PosPC$ or $NegPC$ has a time series property defined by only sign of daily profits. This measure captures the distribution of daily profits without using the magnitude of daily profits, which distinguishes it from volatility or skewness that incorporates magnitude.

2.3.2 Survival Analysis of Investor's Withdrawal

To investigate investor's withdrawal decision from a fund, I use survival analysis to evaluate how long investors in the sample typically stay with a fund before withdrawing. I use a shared frailty model, which extends the Cox proportional hazards model with a random cluster-specific intercept. A key advantage of the frailty approach is that it accounts for unobserved heterogeneity across clusters. In this paper, I use investor-specific cluster as (1) there are repeated observations per each investor-fund pair over time (2) it is reasonable to assume that some investors are habitually more prone to withdraw than others.

In the survival analysis, the primary focus is to model the hazard function, which is the instantaneous rate of occurrence of the event. Here the event of interest $leave_{ijt}$=1 if investor i withdraws from fund j on date t, and $leave_{ijt}$=0 otherwise. The model includes multiple observations per each investor-fund pair, one for every trading day t until the investor's withdrawal from the fund. Correspondingly, the hazard function for the investor i, fund j pair is

$$h_{ij}(t) = h_0(t)\exp(\boldsymbol{\beta}' X_{ijt} + \alpha_i) \tag{2.3}$$

where $h_0(t)$ is the baseline hazard function which represents the hazard when X_{ijt} and α_i are all zero, X_{ijt} is the vector of covariates, $\boldsymbol{\beta}$ is the vector of regression coefficients, and α_i is the random effect associated with the investor i. In terms of the shared frailty ν_i defined by $\nu_i = \exp(\alpha_i)$, the hazard function can be re-written as

$$h_{ij}(t) = h_0(t)\nu_i \exp(\boldsymbol{\beta}' X_{ijt}) \tag{2.4}$$

Assume that the shared frailty ν_i has a gamma distribution, the model is then fitted by the penalized likelihood method.

2.4 Extrapolation and Fund Investment Performance

This section contains three parts. I first show that investors in my sample extrapolate investment performance from past performance consistency in addition to past performance and risk. I then show that past performance or its consistency does not positively predict future performance of fund investment. Lastly, I show that OECD-investors are more prone to extrapolate than non-OECD investors, though neither group of investors profits from extrapolation.

2.4.1 Test for Extrapolation

Assume that the investor forms performance expectation based on his/her own past investment experience with the fund. The extrapolation hypothesis posits that the investor expects high (low) return if past performance is consistently positive (negative). Correspondingly, the hypothesis predicts that an investor will stay with (withdraw from) the fund if recent performance is consistently positive (negative).

To test this hypothesis, I estimate the shared frailty model shown in Equation 2.4. The model includes multiple observations per each follower-fund pair ij, one for every trading day t until the follower i's withdrawal from the fund j. The dependent variable $leave_{ijt}$=1 if follower i withdraws from fund j on date t, and $leave_{ijt}$=0 otherwise. The main independent variables are $PosPC$ and $NegPC$ shown in Equations 2.1 and 2.2, which respectively measure the positive and negative performance consistencies of fund investment j from the perspective of follower i over $t-1$ to $t-10$ trading days (i.e. the past two weeks).

The coefficient on $PosPC$ or $NegPC$ reflects the change in the hazard rate when the fund investment performance is consistently good or poor. In Table 2.2 (a), a negative coefficient on $PosPC$ implies that investors are less likely to withdraw when experiencing recent consistently good performance. A positive coefficient on $NegPC$ implies that investors are more likely to withdraw when experiencing recent consistently poor performance. The coefficients on both $PosPC$ and $NegPC$ are negatively and positively significant, respectively with p-values less than 0.01. This indicates extrapolation.

Table 2.2 (b) adds a set of controls that may correlate with the withdrawal decision. The additional indepdent variable $gain\_all$ equals one if investor's current cumulative net profits from a fund since the investment starting date are positive, otherwise zero; $gain\_10days$ equals one if investor's cumulative net profits from the fund are positive over the past 10 trading days, otherwise zero. These two $gain$ dummy variables are included to account for potential reverse-disposition effect in delegated assets as motivated by literature (Chang, Solomon, and Westerfield, 2016). To control for investment performance and risk, the independent variables $mean$, sd and $skewness$ refer to the mean, standard deviation, and skewness of investor's net profits (in 10,000s dollars) from fund j measured over the past 10 trading days, respectively. I add an interaction term $gain\_10days * sd$ to the model as I observe that the effect of standard deviation on investor withdrawal differs when facing recent gains vs. losses. To control for investor's exposure to the fund investment, the independent variables $trades$ and $position\_size$ are investor's average number of daily open trades and average trading position size (in 10,000s dollars) resulting from the fund investment measured over the past 10 trading days, respectively. The two variables $trades$ and $position\_size$ can also proxy for investor's account size and leverage.

In both specifications Table 2.2 (a) and (b), the coefficient on $PosPC$ or $NegPC$ is qualitatively similar and statistically significant at the 1% level. The results confirm that investors rely on past performance consistency for withdrawal decision. Moreover, the negatively significant coefficient on $mean$ implies that investors are less likely to withdraw when getting higher past performance. This is in line with the notion that investors extrapolate from past performance. Interestingly, in my sample standard deviation of profits appears to have an asymmetric effect on investor withdrawal. The effect of standard deviation on withdrawal is $-1.245+1.509*gain\_10days$. When experiencing losses in the last two weeks or $gain\_10days = 0$, so the effect of standard deviation is -1.245 suggesting that investors are less likely to leave a fund if standard deviation is large. When experiencing gains in the last two weeks or $gain\_10days = 1$, the effect of standard deviation is $-1.245 + 1.509 * 1 = 0.264$ suggesting that investors are more likely to leave a fund if standard deviation is large. This implies that investors tend to tolerate more performance volatility when facing losses, but not so much when facing gains.

2.4.2 Test for Performance Predictability

To test for the relation between measures of past performance and future performance of fund investment, I use OLS panel regression incorporating investor and time fixed effects as follows:

$$profit_{ij,t+1} = \boldsymbol{\beta}' X_{ijt} + \alpha_i + \eta_t + \epsilon_{ijt} \tag{2.5}$$

where $profit_{ij,t+1}$ is the unrealized dollar net profit of follower i's investment in fund j on the next trading day. X_{ijt} denotes the vector of covariates. α_i and η_t denote investor and day-of-the-week fixed effects, respectively.

Table 2.3 reports the regression results from Equation 2.5. The coefficient on either $PosPC$ or $NegPC$ is insignificant and this suggests past performance consistency does not predict future performance. In addition, the coefficients on investment risk measures such as sd and $skewness$ are also insignificant. However, the coefficient on past investment performance $mean$ is significantly negative at 5% level indicating performance reversal. Also $gain\_all$ is significantly negative at 1% level and $gain\_10days$ is weakly significantly positive at 10% level, which provide mixed evidence about the link of investor's reverse-disposition effect to future performance. In addition, the significantly negative coefficients on $trades$ and $position\_size$ indicate investor with higher exposure to the fund investment also has lower future performance.

Combined with earlier finding on extrapolation, the result suggests that the past performance consistency on fund investment has a significant effect on subsequent withdrawal decision, even though it does not help predict the future performance. This is consistent with the over-extrapolation hypothesis, in which investors extrapolate based on past performance sequence and do not benefit by doing so.

2.4.3 OECD vs. Non-OECD Investors

In this subsection, I test extrapolation hypothesis and performance predictability within subsamples of OECD and non-OECD investors respectively. Table 2.4 reports the hazard estimates from Model 2.4 by investor group. The OECD investors have lower coefficient on $PosPC$ and higher coefficient on $NegPC$ than the non-OECD investors. This implies that fund withdrawals of OECD investors are more sensitive to performance consistency. The OECD investors also have lower coefficient on $mean$, which suggests that OECD investors also extrapolate more from past performance. The likelihood-ratio tests confirm that each of the three coefficients on $PosPC$, $NegPC$, and $mean$ is significantly different between the two groups of investors at 1% level[10].

Table 2.5 reports the OLS estimates from Model 2.5 by investor group. For both investor groups, effects of positive and negative performance consistencies as well as standard deviations of performance are not significant. For OECD investors, past performance is not significantly related to future performance, whereas for non-OECD investors past performance is negatively correlated.

The findings suggest that extrapolation influences more on investors from more developed countries, though it does not help with future performance for either group of investors.

2.5 Conclusion

Recent literature has increasingly recognized the role of investor extrapolation in asset allocation and asset pricing. Using novel brokerage data from Zulutrade a social trading platform, I provide empirical evidence that investors over-extrapolate based on past fund performance, and that besides raw performance, past performance consistency is salient to investors when forming expectations.

[10] To compare coefficients between two groups, I conduct likelihood-ratio test for each coefficient of interest respectively. Suppose the coefficient is on $PosPC$. I test goodness-of-fit between two models. Model 1 uses full sample (both OECD and non-OECD investors), all current covariates plus $OECD$, an indicator variable that takes 1 for OECD investor and 0 otherwise. Model 2 adds an interaction term $OECD * PosPC$ to Model 1. Let $\ln(L_1)$ and $\ln(L_2)$ be the max log likelihood of Models 1 and 2, respectively. The test statistic is $2 * (\ln(L_1) - ln(L_2))$, which approximately follows a Chi-Square distribution with 1 degree of freedom

The micro-level data allow me to link fund withdrawal to past performance at individual investor level. I find evidence consistent with extrapolative beliefs, in that investors not only extrapolate past performance, but also are more (less) likely to withdraw when experiencing recent consistent poor (good) performance after controlling for past performance and risk. This finding, combined with the later result that past performance, its consistency and volatility are non-positively related to future performance, suggests that extrapolative fund investing is unprofitable. That is investors over-extrapolate past fund performance. This effect of extrapolation on fund withdrawal is particularly pronounced for OECD-investors in contrast to non-OECD investors.

This study is the first that I am aware of to examine extrapolation for individual fund investors on a global scale. My findings generalize extrapolation behavior in non-delegated assets such as stocks to delegated assets such as mutual funds. I also propose that performance consistency could play a role in future research on extrapolation. The findings are also of interest to regulators. In the U.S., the Securities and Exchange Commission (SEC) currently requires funds to tell investors that a fund's past performance does not necessarily predict future results. In light of the findings in this research, asset management firms should add that a fund's past performance, past performance consistency, and past performance volatility do not necessarily predict future results.

2.6 Tables

Table 2.1: Summary Statistics for Investor-fund Pairs

This table reports summary statistics for positive performance consistency ($PosPC$), negative performance consistency ($NegPC$), reverse disposition effect measures ($gain\_all$ and $gain\_10days$), performance and risk ($mean$, sd and $skewness$), fund exposure measures ($trades$ and $position\_size$) as well as the number of times of re-investment ($reinvest$). The sample is based on the constructed daily panel of investor-fund pair observations from Nov. 2007 to Dec. 2016. $PosPC$ and $NegPC$ are defined in Equations 2.1 and 2.2, which respectively capture the positive and negative consistencies of investor's daily net profits from fund over the past 10 trading days. $gain\_all$ equals one if investor's cumulative net profits from fund since the first investing date are positive on day t, otherwise zero; $gain\_10days$ equals one if investor's cumulative net profits from fund are positive over the past 10 trading days, otherwise zero. $mean$, sd and $skewness$ refer to the mean, standard deviation, and skewness of investor's daily net profits from fund measured over the past 10 trading days, respectively. $trades$ and $position\_size$ are investor's average number of daily open trades and average trading position size from investing in fund measured over the past 10 trading days, respectively. $reinvest$ is the number of times investor has invested in the fund in the past as of day t.

	Mean	Median	Pctl25	Pctl75	SD
PosPC	0.12	0.00	0.00	0.20	0.21
NegPC	0.05	0.00	0.00	0.10	0.17
gain_all	0.53	1.00	0.00	1.00	0.50
gain_10days	0.52	1.00	0.00	1.00	0.50
mean ($)	-1.51	0.10	-2.09	1.97	94.37
sd	47.00	9.43	3.65	27.19	304.65
skewness	0.01	-0.01	-0.74	0.73	1.20
trades	3.41	1.90	1.00	4.00	5.18
position_size ($)	4,201.24	1,286.06	870.95	2,553.68	23,056.94
reinvest	0.16	0	0	0	0.53

Table 2.2: Test for Extrapolation: Hazard Estimates

This table presents estimates of the determinants of the hazard rate to fund withdrawal using the shared frailty model below:

$$h_{ij}(t) = h_0(t)\nu_i \exp(\beta_1 PosPC_{ijt} + \beta_2 NegPC_{ijt} + \boldsymbol{\beta}' X_{ijt})$$

$h_0(t)$ is the baseline hazard function. ν_i is the shared frailty associated with the investor i. The dependent variable $leave$ is one for the (investor, fund, and trading day) triplet if investor i withdraws from fund j on day t, otherwise zero. The main independent variables are $PosPC$ and $NegPC$, which respectively capture the positive and negative consistencies of investor i's daily net profits from fund j over the past 10 trading days (i.e. $t-1$ to $t-10$). The rest of the independent variables are described in Table 2.1, but the variables $mean$, sd, and $position\_size$ here are measured in 10,000s of dollars. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(a)	(b)
PosPC	-1.499^{***}	-0.660^{***}
	(0.015)	(0.018)
NegPC	0.965^{***}	0.503^{***}
	(0.012)	(0.013)
gain_all		-0.523^{***}
		(0.006)
gain_10days		-0.276^{***}
		(0.008)
mean		-3.639^{***}
		(0.414)
gain_10days*sd		1.509^{***}
		(0.341)
sd		-1.245^{***}
		(0.221)
skewness		-0.082^{***}
		(0.003)
trades		-0.006^{***}
		(0.001)
position_size		-0.003
		(0.002)
reinvest		-0.095^{***}
		(0.005)
Observations	7,227,740	7,016,000
Investors	15,073	15,073
Funds	12,086	12,086
Investor-fund pairs	163,774	163,768

Table 2.3: Test for Performance Predictability: OLS Estimates

This table presents results from the following OLS regression:

$$profit_{ij,t+1} = \beta_1 PosPC_{ijt} + \beta_2 NegPC_{ijt} + \boldsymbol{\beta}' X_{ijt} + \alpha_i + \eta_t + \epsilon_{ijt}$$

The dependent variable $profit_{ij,t+1}$ is the unrealized net profit of follower i's investment in fund j on the next trading day. The main independent variables are $PosPC$ and $NegPC$, which respectively capture the positive and negative consistencies of investor i's daily net profits from fund j over the past 10 trading days. The rest of the independent variables are described in Table 2.1. Investor and day-of-the-week fixed effects are α_j and η_t, respectively. Robust standard errors double-clustered by investor and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

PosPC	0.952
	(2.201)
NegPC	1.950
	(2.742)
gain_all	-1.518^{***}
	(0.558)
gain_10days	2.189^{*}
	(1.160)
mean	-0.113^{**}
	(0.054)
sd	-0.007
	(0.021)
skewness	-0.183
	(0.362)
trades	-0.649^{***}
	(0.124)
position_size	-0.0005^{***}
	(0.0001)
Investor FE	Y
Day-of-the-week FE	Y
Observations	6,838,920
Investors	14,880
Funds	11,636
Investor-fund pairs	156,892
R^2	0.004
Adjusted R^2	0.001

Table 2.4: Test for Extrapolation: Hazard Estimates by Investor Group

This table presents estimates of the determinants of the hazard rate to fund withdrawal using the shared frailty model described in Table 2.2. OECD column presents hazard estimates for the subsample consisting of investors from OECD member countries, and non-OECD column presents hazard estimates for the subsample consisting of the rest of investors. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	OECD	non-OECD
PosPC	-0.660^{***}	-0.653^{***}
	(0.023)	(0.032)
NegPC	0.523^{***}	0.451^{***}
	(0.016)	(0.022)
gain_all	-0.544^{***}	-0.484^{***}
	(0.007)	(0.010)
gain_10days	-0.280^{***}	-0.271^{***}
	(0.010)	(0.014)
mean	-4.752^{***}	-3.193^{***}
	(0.655)	(0.622)
gain_10days*sd	1.718^{***}	1.548^{**}
	(0.502)	(0.491)
sd	-1.360^{***}	-1.224^{***}
	(0.304)	(0.344)
skewness	-0.081^{***}	-0.084^{***}
	(0.003)	(0.004)
trades	-0.007^{***}	-0.005^{***}
	(0.001)	(0.001)
position_size	-0.003	-0.006^{*}
	(0.002)	(0.003)
reinvest	-0.088^{***}	-0.107^{***}
	(0.006)	(0.010)
Observations	5,036,891	1,979,109
Investors	8,609	6,464
Funds	9,128	6,967
Investor-fund pairs	108,227	55,541

Table 2.5: Test for Performance Predictability: OLS Estimates by Investor Group

This table presents results from the OLS regression described in Table 2.3. OECD column presents OLS estimates for the subsample consisting of investors from OECD member countries, and non-OECD column presents OLS estimates for the subsample consisting of the rest of investors. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	OECD	non-OECD
PosPC	−0.573	2.490
	(1.959)	(2.958)
NegPC	4.099	−1.016
	(2.681)	(3.218)
gain_all	−1.382***	−2.273*
	(0.470)	(1.203)
gain_10days	1.162	3.666**
	(1.047)	(1.592)
mean	−0.052	−0.155***
	(0.040)	(0.055)
sd	−0.012	−0.005
	(0.027)	(0.020)
skewness	−0.283	−0.184
	(0.316)	(0.509)
trades	−0.658***	−0.564***
	(0.206)	(0.170)
position_size	−0.0004***	−0.001**
	(0.0001)	(0.0003)
Investor FE	Y	Y
Day-of-the-week FE	Y	Y
Observations	4,919,346	1,919,574
Investors	8,520	6,360
Funds	8,798	6,675
Investor-fund pairs	104,030	52,862
R^2	0.003	0.004
Adjusted R^2	0.002	0.001

CHAPTER 3

MEDIA SENTIMENT IN A MOBILE ERA

3.1 Introduction

There are a number of studies that document the predictive value of media sentiment on stock market reactions[1]. However, one aspect for this influence is less explored: the question of how investors get the news. A 2017 survey conducted by the Pew Research Center finds that 85% of U.S. adults get their news on a mobile device at least some of the time with higher usage among younger groups of adults, and nearly two-thirds of those who get news on both mobile and desktop prefer mobile[2]. However, attention to news is significantly different across mobile and desktop users (Dunaway et al., 2018). As most mobile usage is done on the go, people reading news on mobile spend less than one-third as much time compared to when reading news on desktop[3]. Research also suggests that information is harder to process due to the smaller screens of mobile devices (Kim and Sundar, 2016). Overall, mobile users are mostly characterized as young with limited attention. Their reaction to financial news provides an interesting setting to test theories of media sentiment.

Motivated by the increasingly mobile trend in digital news consumption, this paper proposes novel measures of media sentiment based on Google mobile and desktop search results, and posits that mobile users are less informed

[1] For example, Barber and Loeffler (1993); Huberman and Regev (2001); Tetlock (2007); Engelberg and Parsons (2011); García (2012); Chen et al. (2014)

[2] See more at `http://www.pewresearch.org/fact-tank/2017/10/04/key-trends-in-social-and-digital-news-media/`

[3] The differences in mobile and desktop users' browsing habits come from a report from Taboola, an online marketing company. See more at `https://blog.taboola.com/desktop-vs-mobile-getting-traffic-want/`

than desktop users, therefore more likely to move prices away from fundamentals affecting future stock returns and liquidity. In particular, I construct mobile and desktop media sentiments respectively using unique web scraped data, which consist of actual Google search results on both mobile and desktop platforms. My choice of Google is motivated not only by Google's popularity as the largest search engine but also by its increased focus to optimize mobile search as most people today search Google from mobile[4]. Interestingly, when searching the same term about a stock on mobile and desktop, Google sometimes returns drastically different search result contents. This is because Google uses a different searching algorithm on mobile devices that (1) prioritizes fast and mobile-friendly website[5] (2) crawls content from the mobile version of a website which sometimes however carries content different from its desktop version (e.g. see the differences in the top three results in Figure 3.1 and Figure 3.2).

According to the Pew Research Center, about 20% of the time online news consumers get their news from search engines while most of the time (about 36%) they get it directly from a news website. I posit that views expressed about stocks from Google search have the largest impact on naive individual or retail investors as opposed to sophisticated traders who have access to better sources of information. Given that mobile users are generally younger, they are likely to be more overconfident (Barber and Odean, 2001). They are also more likely to be distracted making decisions based on limited absorbed information. As such, I expect sentiment of mobile users is better correlated with stock market reactions than that of desktop users. In this paper, sentiment refers to the level

[4]See more at `https://techcrunch.com/2017/12/20/googles-mobile-first-search-index-has-rolled-out-to-a-handful-of-sites/`

[5]In February 2016, Google began to favor AMP sites in ranking mobile search results. AMP-enabled sites use Google's Accelerated Mobile Page technology to load up pages significantly faster than regular mobile pages.

of beliefs of irrational traders or noise traders relative to that of rational traders (De Long et al., 1990a). Following the empirical evidence reported in Tetlock (2007), I expect negative sentiment predicts return reversal at longer horizons, as noise trading or trading based on non-information-based reasons (e.g. liquidity needs) should affect short-term returns that will be reversed in the long run. Also, models of noise traders predict that unusually high or low values of sentiment will generate high volume from noise trading (De Long et al., 1990a; Campbell, Grossman, and Wang, 1993). In the models of Glosten and Milgrom (1985) and Admati and Pfleiderer (1988), increased noise trading decreases the market maker's adverse selection cost leading to increased liquidity reflected by lower spreads and more depth; Bloomfield, O'Hara, and Saar (2009) presents experimental evidence in support of this view. Consequently, I hypothesize that at individual stock level, high negative sentiment will forecast higher abnormal returns, and large absolute differences in negative sentiment will forecast higher liquidity with lower spread and higher volume, and the effects should be stronger from mobile users.

In my empirical work, to quantitatively measure the interactions between Google search results and stock market, I first conduct textual analysis using the text content of stock-specific search results for each stock in the S&P 500 index. The mobile and desktop media sentiments of a stock are computed by using its mobile and PC desktop results, respectively. The text content consists of the top 30 search results about the stock (equivalent to first 3 pages of results). The search term is the stock ticker plus the word "stock" (e.g. "MCHP stock" for Microchip Technology) as motivated by literature (Da, Engelberg, and Gao, 2011). On each calendar day after 9:30 a.m. New York time, the search is performed on mobile and desktop platforms, respectively. My sample ranges from late August 2017 to December 2018 covering 330 trading days. In contrast to the traditional sentiment analysis methods in prior literature (Tetlock, 2007; Loughran and McDonald, 2011), this paper harnesses the advantages of an open source Python package VADER (Valence Aware Dictionary and sEntiment Reasoner), a popular lexicon and rule-based natural language processing tool specifically attuned to sentiments expressed in microblog-like contexts. Compared to the traditional method that calculates the fraction of words classified as positive or negative based on a fixed dictionary, VADER provides some key improvements: (a) capture both the polarity (positive or negative) as well as the intensity (e.g. "good" vs. "great"). (b) use syntactical cues to handle negation (e.g. "not good"). Moreover, unlike the data sources of prior literature which are mostly financial in nature, Google are more microblog-like indexing short truncated content from a variety of websites including blogs, forums, social and financial news media. Hence VADER is well-suited to have a precise gauge of sentiment in such context.

I then estimate the intertemporal links between mobile or desktop sentiment and stock market building on Tetlock (2007) who uses vector autoregressions (VARs). In this paper, the main variables for measuring stock market activity are abnormal returns, effective bid-ask spread and volume (turnover ratio). Different from Tetlock (2007)'s focus on a time series - Dow Jones index, I have a panel data structure as I observe multiple stocks each day and over time. For this reason, I adopt panel VARs method for my analysis which also captures time-invariant unobserved heterogeneity across both firms and days of the week. For estimation of the VAR model, I use equation-by-equation OLS which is equivalent to Granger causality test (Granger, 1969). An advantage of this approach is that I'm able to examine the mutual Granger causality or one-directional Granger causality between sentiment and different indicators of market activity.

This paper's key findings are that high levels of negative mobile or desktop sentiment predict higher abnormal returns in the following week with mobile serving as a more significant predictor than desktop; for liquidity, high absolute differences in negative sentiment have a significantly negative impact on the next trading day's spread with mobile being more significant than desktop; however, there is suggestive evidence that only mobile sentiment has impact on volume. These findings suggest that views expressed in Google search results reflect investor sentiment. Furthermore, I show that the return reversal prediction from negative sentiment is mainly driven by stock over-pricing rather than under-pricing. More specifically, as investors become less pessimistic or more optimistic, the stock is over-priced and will later revert back to its fundamental value (i.e. lower future returns), whereas as investors become more pessimistic, the stock is unlikely to be under-priced and have reversal. This supports the view that short sale constraints limit the ability of rational traders to correct overpricing. In addition, the effect of mobile sentiment on returns becomes more pronounced than desktop sentiment in stocks of high retail interests. The findings also provide suggestive evidence that mobile users help supply liquidity resulting in lower spreads on the next trading day. To compare the respective roles of mobile and desktop, I include lags of mobile and desktop sentiments together for prediction, the coefficients on the lags of mobile sentiment are often more significant. Clearly, mobile sentiment serves as a more prominent predictor for market reactions than desktop. This supports the idea that mobile users are more susceptible to the influence of media sentiment. One potential explanation is that the young naive demographic combined with short attention span associated with mobile use result in less informed trading. Such interpretation points to the growing importance of addressing mobile media when evaluating the effect of media sentiment.

To examine if my sentiment measures reasonably proxy for investor sentiment, I use past market behavior to predict the values of my sentiment measures. I find that negative abnormal returns predict more negative sentiment, which implies positive feedback trading which seeks to capitalize on an initial price move by buying when price rises and selling when price falls (De Long et al., 1990b). Also high spread predicts more negative sentiment, which supports the notion that investors dislike illiquidity. In addition, high trading volume increases negative sentiment, which could suggest that high volume stocks exhibit glamor characteristics with lower future returns leading to more negative future sentiment (Lee and Swaminathan, 2000).

The paper most similar to mine is Joseph, Babajide Wintoki, and Zhang (2011), who study the market abnormal returns and volume forecasting ability of the investor sentiment using public search volume index (SVI) data from Google Trends. My study differs in the type of data and methodologies. I use unique data of actual observed Google search results which differentiate between mobile and desktop platforms. I expect the actual content of search results reveals additional information because the tone of results can lead to disagreement in beliefs which in turn leads to trading. Also the separate treatment of mobile and desktop contents is novel to my study as I demonstrate mobile plays a more significant role in the domain of financial markets nowadays.

Moreover, the paper provides suggestive empirical evidence that links liquidity to sentiment. A related study is Liu (2015) who uses survey-based investor sentiment index and shows that high sentiment increases market liquidity. Unlike the market-wide sentiment index and liquidity in Liu (2015), I measure sentiment at the individual stock level, and in addition to turnover, I construct a high quality stock-level liquidity measure using the national best bid and offer (NBBO) prices and trades from NYSE Daily Trade and Quote (DTAQ) data. This allows me to examine how sentiment interacts with liquidity more granularly over time.

The remainder of this paper is organized as follows. Section 3.2 presents the data, sentiment analysis method, stock market variables and summary statistics. Section 3.3 examines the relation between sentiment and stock return. Section 3.4 examines the relation between sentiment and stock liquidity. Section 3.5 concludes the paper.

3.2 Data

This study uses data collected from stock ticker search results from Google, financial data from Compustat and the Center for Research in Security Prices (CRSP), as well as daily trade and quote data from NYSE DTAQ. The sample period is from August 2017 to December 2018.

3.2.1 Google search

I used a program to automatically search the tickers of all stocks in the S&P 500 index on Google and extract the text content of search results every day. The search term is the stock ticker plus the word "stock", for example for Microchip

Technology, it's "MCHP stock." Since Google search algorithm treats mobile and personal computer (PC) users differently[6], the same term is searched on both android phone and desktop PC platforms respectively. To control for influence of user's location on search results, the location on both platforms is set to be Wall Street in New York City.

For each stock ticker, I focus on three sections on its search results page to form the text corpus: the results section which includes anchor text (i.e. link title) and snippet of the results, the news section (if available) which includes text of stock-specific news headlines, the related searches section (if available) which includes related search terms. For the results section, a conventional Google page shows 10 results at default, I increase the limit to 30 results which amount to pooling results from three conventional pages. I extract the anchor texts and snippets from all the results to form the text corpus. For the news section, sometimes Google may display headlines of news or top stories related to the stock ticker at the top of the results page; the news are stock-specific and are typically sourced from financial news websites. For the related searches section, Google often display keyword suggestions that are relevant to the searched stock ticker at the bottom of the results page. Overall, the text content of the results section makes up the main text corpus, which is combined with news headlines and related searches if any of those sections is available.

On each trading day, each stock has two platform-dependent sentiment measures constructed separately from the text corpus of mobile search results and of PC search results on that day. For non-consecutive trading days, when the market was not open, I first find the previous trading day and then assign it the average of the daily sentiment measures since that day until the market

[6]See footnote 5

next opens. The platform induced variation in Google results is stressed here as sometimes it leads to notably different sentiment scores. For instance, "MCHP stock" (i.e. Microchip Technology stock) is associated with its mobile sentiment scores in Figure 3.1 and its PC sentiment scores in Figure 3.2 on March 6th, 2018. In this example, PC results have lower negative intensity and higher positive intensity than the corresponding mobile results. This is partly because PC results show the positive phrase "designed with customer innovation in mind" repeatedly in the result snippets ranked numbers 2 and 3. In contrast, such phrase does not appear in mobile results.

3.2.2 Sentiment analysis using VADER

In natural language processing, there are two major approaches to conduct sentiment analysis: machine learning methods and lexicon-based methods (Guo, Shi, and Tu, 2016). The machine learning (ML) methods often use supervised classification framework training classifiers with labeled data. For instance, Antweiler and Frank (2004) start by manually classifying 1,000 messages from stock message boards to train a Naive Bayes classifier which is then implemented out-of-sample. In textual analysis the same word could change meaning in different contexts, the ML methods work well when the training data have researchers' desired context (e.g. train with blog posts labeled with good, neutral or bad to classify unlabled blog posts, train with Twitter stock tweets labeled with buy, hold or sell to classify unlabeled stock tweets). Depending on the context of interest, training data might not be readily available and manual classification would be then required to first construct the training data. Unlike the ML methods, the lexicon methods don't need to train but instead use predefined lists of words, in which each word is associated with a specific sentiment. The lexicon methods are quite common for evaluating sentiment in prior literature. Tetlock (2007) uses the Harvard IV-4 psychosocial dictionary. Loughran and McDonald (2011) have compiled positive and negative word lists designed for use in financial context. Prior literature often follows a lexicon-based textual analysis that calculates the fraction of words classified as positive or negative based on a fixed dictionary. However, there are some drawbacks with this simple technique such as not capturing the intensity of positivity or negativity (e.g. "good" vs. "great", "bad" vs. "very bad") and unable to handle negation (e.g. "not good" or "not bad").

To have a more precise gauge of sentiment, this paper uses VADER (Valence Aware Dictionary and sEntiment Reasoner) which is a popular lexicon-based method combined with five rules based on grammatical and syntactical cues to convey changes to sentiment intensity (Hutto and Gilbert, 2014). The sentiment lexicon used in VADER is sensitive to both the polarity and the intensity of sentiments (e.g. "great" has higher positive ratings than "good"). The lexicon is also gold-standard quality attuned especially to microblog-like contexts and has been validated by humans. The five rules used in VADER include treatments for: (1) punctuation (e.g. intensity is increasing in the number of exclamation points "!"s); (2) capitalization (e.g. "GOOD" is more intense than "good"); (3) degree modifiers (e.g. "very good" is more intense than "good"); (4) contrastive conjunction "but" to shift the polarity (e.g. "American Airlines Group holds several positive signals, but we still don't find these to be enough for a buy-recommendation" has mixed sentiment, with the latter half after "but" dictating the overall rating.); (5) tri-gram examination to identify negation (e.g. "this company is not effective at turning revenues into bottom line profit" where in the tri-gram "is not effective", "is not" which precedes the positive word "effective" signals negation and the polarity is flipped into negative).

The paper makes use of VADER python library computing multidimensional sentiment scores for a given text corpus: positive, negative and neutral. The three scores are ratios for proportions of text that fall in each category. As mentioned earlier, the simple lexicon method divides total number of positive, negative or neutral words by the total number of words to compute positive, negative, or neutral scores. VADER computes the scores differently dividing the total magnitude of positive, negative, or neutral intensity by the total magnitude of intensity. Just like the scores from the simple method, VADER's scores range from 0 to 1 with the three add up to be 1.

As literature on textual analysis is growing, it has become increasingly important to choose a transparent, replicable and suitable method when translating text into quantitative measures. A recent survey paper Loughran and McDonald (2016) points out "the best way to avoid the numerous and substantial tripwires in textual analysis is to carefully consider the ability of a program, word list, and statistical method to work effectively in the specific context of application." Since VADER appears to be a nascent method to finance literature, this study justifies using VADER for two main reasons. (1) VADER lexicon performs exceptionally well in the social media context comparable to individual human raters and also adapts well in diverse contexts (Hutto and Gilbert, 2014). A recent benchmark study on 24 sentiment analysis methods (lexicon or ML-based) find VADER performs the best in 3-class classification (i.e. positive, negative and neutral) using testing data including Twitter tweets, BBC and New York Times comments (Ribeiro et al., 2016). In this paper, the text content of search results are compiled by Google from a variety of websites not limiting to financial news media which are the general focus of prior literature. Thus, I consider social media to be a suitable context, and posit that VADER is appropriate for this study. (2) unlike ML methods, VADER is purely lexicon-based, non-black-box, easily replicable and more computationally efficient. It also does not require training data which could be a labor intensive to acquire or construct and sometimes error prone process (e.g. due to lack of representative word features in the training data).

3.2.3 Abnormal returns, effective spread and volume

The abnormal returns are defined as the difference between raw returns and returns on a value-weighted portfolio of firms with similar size, book-to-market ratio, and past returns (Daniel et al., 1997). In particular, I follow the appendix in Daniel et al. (1997) to construct 125 portfolios based on a conditional triple-sort first on quintiles of every U.S. firm's market equity value, then further on quintiles of book-to-market ratio, and finally on quintiles of momentum; the portfolios are re-sorted every year in July. The market value and firm financial data are from CRSP and Compustat databases. To calculate the abnormal return of a stock in my sample on date t, I first find its matched portfolio and subtract the date t value-weighted return of the portfolio from the stock's raw return.

To construct the effective spread, I use DTAQ data, which include all trades and quotes for stocks traded on NYSE, Nasdaq and the regional exchanges. For a given stock-day, I follow Holden and Jacobsen (2014) to construct NBBO prices and clean the trades. I match a given trade to the NBBO that was in effect one millisecond earlier (i.e. a one millisecond lag). I use Lee and Ready (1991) convention to determine if the trade is a buy or a sell. Then I calculate the dollar effective spread for the trade, which is $2D(P-M)$ where D is a dummy variable that equals +1 if the trade is a buy and -1 if the trade is a sell, P is the traded price, and M is the midpoint of the NBBO quotes matched to the trade. Lastly, I average across all the trades to get the daily effective spread for the stock. As the level of spread is not stationary, I focus on log difference of spread which approximates the percentage change in spread in regressions.

For volume measure, I use log value of stock daily turnover ratio and follow a detrending methodology based on Campbell, Grossman, and Wang (1993) because the level of log turnover is not stationary. Specifically for a given stock on date t, I first compute the volume trend as a rolling average of its past 30 days of log turnover, and then subtract the volume trend from date t's log turnover.

3.2.4 Summary statistics

The sample encompasses 379 U.S.-based S&P 500 stocks covering 330 trading days from August 2017 to December 2018[7]. There are 116,254 stock-trading day observations.

Table 3.1.A shows summary statistics for sentiment scores computed separately from mobile and PC search results. The average negative sentiment score for mobile is 1.83%; and the average positive score for mobile is 6.20%. In comparison, the average negative sentiment score for PC is 1.79%; and the average positive score for PC is 6.15%. The correlation between negative mobile and

[7]Stock selection follows Campbell et al. (1997), in which U.S.-based stocks are identified as the subset of securities that have a share code (SHRCD field in CRSP) value of either 10 or 11; the stocks in the sample must also have sufficient information to compute book-to-market ratio for the fiscal year.

negative PC sentiment scores is 0.922. The correlation between positive mobile and positive PC sentiment scores is 0.951. Despite the high correlations, I find sentiment difference between the two platforms, either negative or positive, is significantly different from 0 at the 5% level. That is on average for a given stock, mobile search results have higher negative intensity and higher positive intensity than corresponding PC results.

Table 3.1.B shows summary statistics for the main variables used in this study for measuring stock market activity. On average for a stock, abnormal returns are slightly negative, effective spread is about 3 cents and daily turnover ratio is 0.9%.

3.3 Sentiment and Returns

I run the main regressions to test whether the sentiment predicts future returns at the individual stock level. Following the literature, I focus on the negative sentiment[8] (Tetlock, 2007; Loughran and McDonald, 2011). I adopt a panel vector autoregressive (Panel VAR) framework in which I simultaneously estimate the relationship between returns and the negative sentiment. All VAR estimates include all lags up to 5 trading days prior to market activity. The endogenous variables in the VAR are individual stock's abnormal returns ($ARet$), negative mobile sentiment score (neg_{mob}) and/or negative PC sentiment score (neg_{pc}). I define a lag operator $L5$ to transform any variable x_t into a row vector consisting of the five lags of x_t (i.e. $L5(x_t) = [x_{t-1}\, x_{t-2}\, x_{t-3}\, x_{t-4}\, x_{t-5}]$).

The VAR model can be expressed as a system of equations estimated

[8] In separate tests, I observe insignificant predictability for positive sentiment.

equation-by-equation using standard ordinary least squares (OLS). Such estimates are consistent without the need to assume the error terms are uncorrelated between the equations. Also, this equation-by-equation OLS has the same setup as testing for Granger causality (Granger, 1969). The main equation of interest is the returns equation in the VAR:

$$ARet_{i,t} = \boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\gamma} L5(ARet_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t} \quad (3.1)$$

where the main independent variables of interest $L5(neg_{\text{mob}})$ and $L5(neg_{\text{mob}})$ are five lags of negative mobile and PC sentiments, respectively. α_i and α_t capture the firm fixed effects and time fixed effects (using day-of-the-week dummies), respectively. T-statistics are computed using robust standard errors double-clustered by firm and day to account for serial- and cross-correlation, as well as heteroscedasticity.

Table 3.2.A reports the regression results from Equation 3.1. It runs a horse race between mobile and PC sentiments. In specification (1.a), I first use five lags of negative sentiment scores constructed from mobile results as the main independent variables. The p-value for the null hypothesis that the five lags of the negative mobile sentiment do not forecast returns is below 0.001 strongly implying that negative mobile sentiment is associated in someway with future returns. The sum of the coefficients on the five lags is 3.1 basis points with p-value below 0.001, which is significantly different from zero[9]. That is for a given stock, a 1% increase in its negative mobile sentiment increases its next week's returns by 3.1 basis points. This suggests that negative mobile sentiment exerts a statistically and economically significant positive influence on the returns of the ensuing trading week. Then in specification (1.b), I use five lags

[9] One basis point equals a daily return of 0.01%.

of negative sentiment scores constructed from PC results as the main independent variables. The table shows qualitatively similar results to the mobile case. Finally in the main specification (1.c), I use both mobile and PC lags as the main independent variables. Together all the mobile and PC lags are jointly significant in forecasting returns with p-value below 0.01. However, in this specification, only the five mobile lags are jointly significant with p-value below 0.1 whereas the five PC lags are not jointly significant with p-value of 0.713. Also, only mobile negative sentiment by itself exerts a statistically significant 2.1 basis points positive influence on the next week's returns with p-value below 0.05. This suggests that mobile content has more significant effect on future returns than PC content, which could be attributed to more noise trading driven by mobile search. Overall, this shows consistent evidence of return reversal caused by negative investor sentiment, especially from mobile users.

To compare the temporal effects of low and high negative sentiments, I sort stocks into two subsamples by either mobile or PC negative sentiment. That is on each trading day, I sort stocks into low (high) negative sentiment group if they have less (higher) than median negative sentiment score. Figure 3.3 plots cumulative abnormal returns of low and high negative sentiment stocks. For low negative sentiment stocks, returns get higher prior to the low negative sentiment but then reverse in the future. Interestingly, for high negative sentiment stocks, returns get lower prior to the high negative sentiment but don't appear to reverse in the future. These findings suggest that the prior return reversal prediction from negative sentiment is mainly driven by stock over-pricing. This is consistent with the concept that short sale constraints limit the ability of rational traders to correct overpricing, but not underpricing (Miller, 1977).

3.3.1 Return Predictability in Subsamples

I now investigate subsamples of stocks in which the effect of mobile sentiment should be more prominent with higher retail interests. I first sort stocks into two subsamples by institutional ownership. I use the institutional holdings data from Factset Lionshares, which collects the mandatory quarterly 13F filings with the SEC. I define institutional ownership as the sum of the holdings of all institutions as a percentage of market capitalization for a stock. At start of each quarter, I sort stocks into low (high) institutional ownership group if they have less (higher) than median institutional ownership in the $t-1$ quarter. Due to the availability of Lionshares data, the subsamples cover from August 2017 to June 2018. I then sort stocks into two subsamples by size. On each trading day, I sort stocks into small (large) size group if they have less (higher) than median market cap in the $t-1$ day.

Table 3.2.B shows OLS estimates of Model 3.1 for each subsample. The five lags of the negative mobile sentiment are jointly significant with p-value of 0.003 in low institutional ownership subsample, and become insignificant in high institutional ownership subsample. In contrast, the five lags of the negative PC sentiment are only weakly jointly significant in high institutional ownership subsample with p-value of 0.054. This implies that the returns for stocks with low institutional ownership are more influenced by mobile sentiment rather than desktop sentiment. Similarly, the five lags of the negative mobile sentiment are jointly significant in small size subsample, and become insignificant in large size subsample. In contrast, the five lags of the negative PC sentiment are not significant in either subsample sorted by size. This implies that the returns for stocks with small cap are more influenced by mobile sentiment rather

than desktop sentiment. Overall, the results suggest that effect of mobile sentiment on returns is stronger among stocks with low institutional ownerships and small sizes.

3.3.2 Predict Sentiment

To see if the sentiment scores provide a reasonable measure of information about stocks from Google search, I use past abnormal returns to predict the negative sentiment scores constructed from Google mobile and PC results, respectively. The VAR equation below describes this relationship for respective platform:

$$neg_{\text{platform}} = \boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\gamma} L5(ARet_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t} \quad (3.2)$$

Table 3.3 presents OLS estimates of γ in the first five rows, which represents the impact of past stock abnormal returns on negative sentiment. For either mobile or PC, Table 3.3 reverses the causal link posited in Table 3.2.A. The table indicates that higher (lower) abnormal returns predict lower (higher) negative sentiment which implies positive feedback trading (De Long et al., 1990b). The sum of the lagged return coefficients for mobile case - specification (2.a) implies that a 1% increase in the prior day's abnormal returns leads to a significant 0.079% decrease in negative mobile sentiment next week with p-value below 0.001. In specification (2.b), similarly for PC sentiment higher abnormal returns predict a significant 0.058% decrease with p-value below 0.001.

3.4 Sentiment and Liquidity

3.4.1 Effective Spreads

It is also interesting to look at the effect of sentiment on stock liquidity. I start the analysis first with dollar effective spreads, which reflect stock illiquidity. Note that theories predict that unusually low or high (i.e. large absolute difference of) negative sentiment will increase noise trading volume (De Long et al., 1990a; Campbell, Grossman, and Wang, 1993). Accordingly, I add five lags of absolute difference between current sentiment and mean of daily sentiments over the past trading week. The absolute difference is denoted by the AD prefix, which is applied to mobile and PC negative sentiments, respectively. To control for directions and past trend of sentiment, I also include five lags of sentiment in the VAR equation below which models the log difference of dollar effective spreads:

$$\begin{aligned} \Delta ESpd_{i,t} = &\boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\theta}_{\text{mob}} L5(ADneg_{\text{mob}}) + \\ &\boldsymbol{\theta}_{\text{pc}} L5(ADneg_{\text{pc}}) + \boldsymbol{\gamma} L5(\Delta ESpd_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t} \end{aligned} \tag{3.3}$$

Table 3.4 presents coefficients $\boldsymbol{\theta}_{\text{mob}}$ (rows 6 to 10) and $\boldsymbol{\theta}_{\text{pc}}$ (rows 16 to 20) on the absolute difference of mobile and PC sentiments, respectively. On a given platform, each $ADneg$ coefficient measures the impact of a 1% increase in absolute difference of negative sentiment on dollar effective spreads in percentages. The results suggest that mobile sentiment plays a more direct role than PC sentiment in forecasting illiquidity. In all specifications (3.a) to (3.c), the five mobile or PC lags of absolute difference of sentiment are jointly significant. However, mobile sentiment appears to have more significant immediate effect on spread:

in specification (3.a), a 1% increase in absolute difference of mobile negative sentiment significantly decreases the next trading day's spread by 38.1 basis points; in specification (3.c), it decreases the next trading day's spread by 30.1 basis points though the significance is weaker. PC sentiment has weaker immediate impact on spreads as the first lag of absolute difference of PC negative sentiment is weakly significant in (3.b) and insignificant in specification (3.c).

This suggests that high absolute difference of negative mobile sentiment may increase future stock liquidity. The prominence of mobile in the results is consistent with the theory of De Long et al. (1990a), as mobile users could be less informed thus help supply liquidity by noise trading.

To see if liquidity affects sentiment, I use past dollar effective spreads to predict the negative sentiment. The VAR equation below describes this relationship:

$$neg_{\text{platform}} = \boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\gamma} L5(\Delta ESpd_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t} \quad (3.4)$$

Table 3.5 presents OLS estimates of γ in the first five rows, which represents the impact of past stock illiquidity on negative sentiment. For either mobile or PC, the table indicates that higher illiquidity predict higher negative sentiment suggesting investor's dislike for illiquid stock. The sum of the five lagged $\Delta ESpd$ coefficients implies that a 1% increase in the prior day's dollar effective spread leads to a significant 0.004% (0.004%) increase in negative mobile (PC) sentiment next week with p-value below 0.01.

3.4.2 Volume

Next, I consider the effect of the negative sentiment on stock volume which also reflects liquidity. The VAR equation below models stock volume measured by detrended log turnover:

$$
\begin{aligned}
Vlm_{i,t} =& \boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\theta}_{\text{mob}} L5(ADneg_{\text{mob}}) + \boldsymbol{\theta}_{\text{pc}} L5(ADneg_{\text{pc}}) + \\
& \boldsymbol{\gamma} L5(Vlm_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t}
\end{aligned}
\tag{3.5}
$$

As shown in the specifications (5.a) and (5.b) in Table 3.6, neither mobile (rows 6 to 10) nor PC negative sentiment (rows 16 to 20) induces a significant immediate increase in future stock volume. Nevertheless, in the specification (5.c), the weakly significant first lag of absolute difference of mobile negative sentiment suggests that high absolute difference of negative mobile sentiment may increase the next trading day's volume. This adds to the earlier results from effective spreads prediction that mobile sentiment relates more to liquidity than desktop sentiment.

Next, it is possible that the effect goes the other way around, namely high trading volume increases negative sentiment. The VAR equation below describes this relationship:

$$
neg_{\text{platform}} = \boldsymbol{\beta}_{\text{mob}} L5(neg_{\text{mob}}) + \boldsymbol{\beta}_{\text{pc}} L5(neg_{\text{pc}}) + \boldsymbol{\gamma} L5(Vlm_{i,t}) + \alpha_i + \alpha_t + \epsilon_{i,t} \tag{3.6}
$$

The results in Table 3.7 suggest that high past trading volume directly forecasts more negative sentiment. For either mobile or PC, the sum of the five lagged volume coefficients is significantly positive with p-value below 0.001. Notice that the strongest and most significant increase in negative sentiment comes from the first lag or the previous day's volume. An interpretation of this finding is that negative sentiment increases as high volume stocks act like glamor stocks, which have lower future returns (Lee and Swaminathan, 2000).

From the perspective of Granger causality in my VAR framework, it seems that sentiment is mutually Granger-causal with either return, effective spread or volume as I find Granger causality in both directions.

3.5 Conclusion

In today's digital environment, using mobile devices to get information has become increasingly popular especially among young people. Compared to those who read news on a desktop, people who read news on a mobile device are less focused on digesting the information. Motivated by the differential user base and technical functionality of mobile and desktop, this paper examines the extent to which mobile and desktop news forecast future market returns and liquidity.

I use unique scraped stock ticker search results data from Google, which sometimes produces drastically different results on mobile and desktop platforms due to its mobile-specific search optimization. I use VADER package to quantify mobile and desktop sentiments, which are then used to evaluate the respective roles of the mobile and desktop contents on the stock market. I find that high levels of negative sentiment predict higher abnormal returns in the following week, and mobile is a more significant predictor than desktop. This return reversal prediction from sentiment is mainly driven by stock over-pricing, and the effect of mobile as opposed to PC sentiment is especially pronounced in stocks of high retail interests. Moreover, mobile users relative to desktop users may help supply more liquidity when the negative sentiment is unusually low or high.

My study distinguishes itself from the prior studies through its focus on differentiating the role of mobile and desktop news. The current gradual shift to mobile news consumption signifies an important change in getting and digesting financial information. Plagued by smaller screens and choppy connection, mobile devices are less conducive to news consumption. Hence there can be a knowledge gap between investors who mostly get their news from mobile and those who mostly get it from desktop. As a result, the less-informed mobile users are more easily swayed by media sentiment to trade. My findings point to the importance of mobile media in disseminating financial news as well as the limits to information-seeking on mobile devices.

3.6 Figures

News

1. What Microchip Technology's $10.15 billion Microsemi deal means for the company, semiconductor industry
2. Mizuho adjusts Microchip, Microsemi after acquisition news
3. Microchip Technology in talks to buy Microsemi: source
4. Microchip to buy Microsemi for about $8.35 billion

Results

1. *MCHP Stock Price - Microchip Technology Inc. Stock Quote (U.S.: Nasdaq) - MarketWatch*
   Microchip Technology Inc. stock price, stock quotes and financial overviews from MarketWatch.
2. *MCHP : Summary for Microchip Technology Incorporat - Yahoo Finance*
   View the basic MCHP stock chart on Yahoo Finance. Change the date range, chart type and compare Microchip Technology Incorporat against other companies.
3. *Microchip Technology - MCHP - Stock Price & News | The Motley Fool*
   Real time Microchip Technology (MCHP) stock price quote, stock graph, news & analysis.
4. *Microchip Technology Inc. - NASDAQ:MCHP - Stock Quote & News - TheStreet*
   View detailed financial information, real-time news, videos, quotes and analysis on Microchip Technology Inc. (NASDAQ: MCHP). Explore commentary on Microchip Technology Inc. and hear what the experts at ...
5. *MCHP Stock Quote - Microchip Technology Incorporated Common Stock Price - Nasdaq*
   Stock quote for Microchip Technology Incorporated Common Stock Common Stock (MCHP) with real-time last sale and extended hours stock prices, company news, charts, and research at Nasdaq.
6. *Microchip Technology Incorporated - MCHP - Stock Price Today - Zacks*
   View Microchip Technology Incorporated MCHP investment & stock information. Get the latest Microchip Technology Incorporated MCHP detailed stock quotes, stock data, Real-Time ECN, charts, stats and ...
7. *MCHP Stock Price & News - Microchip Technology Inc. - Wall Street Journal*
   Microchip Technology Inc. Stock - MCHP news, historical stock charts, analyst ratings, financials, and today's Microchip Technology Inc. stock price.
8. *MCHP - Microchip Technology Inc Stock quote - CNNMoney.com*
   Stocks Fall in Premarket Trade. Mar 02 / GuruFocus News - Paid Partner Content. Dow Tumbles Amid Fears of Global Trade War; S&P 500 and Nasdaq Finish Higher. Mar 02 / TheStreet.com - Paid Partner Content ...
9. *MCHP Microchip Technology Inc. stock and investing information on StockTwits (MCHP)*
   $ MCHP Posting a solid gain today in spite of large mkt sell-off, supporting portfolio holdings crowdthnk.com/ stocks / MCHP · MariaC82Mar 2nd, 10:03 am. Upgrades, Downgrades, Initiations & More: ...
10. *Microchip Tech Stock Quote: MCHP Stock News, Quotes, Analysis | Investors.com - Research - Investor's Business Daily*
    Microchip Tech (MCHP) - MANUFACTURES DIGITAL SIGNAL MICROCONTROLLERS, FLASH MEMORY PRODUCTS, AND MIXED SIGNAL AND INTERFACE DEVICES. GROUP LEADERSHIP. Elec-Semiconductor Mfg Group. MCHP.
11. *MCHP 94.47 +3.48 (Microchip Technology Inc) - Morningstar*
    Morningstar real-time stock quotes, graphs, and independent analysis for MCHP keep you informed. Stay up to date with Microchip Technology Inc stock price and analyst ratings on Morningstar.com.
12. *MCHP: 94.47 0.00 (0.00%) - Stock Quote and News - CNBC*
    Get Microchip Technology Inc (MCHP :NASDAQ) real-time stock quotes, news and financial information from CNBC.
13. *MCHP Stock Price & News - Microchip Technology Inc. - Barron's*
    View the latest MCHP stock price with Barron's. Including historical share prices, analysis, earnings, cash flow and market valuation for Microchip Technology Inc.
14. *MCHP Analysis & News - Microchip Technology Incorporated | Seeking Alpha*
    Microchip Technologies: High Risk, Low Reward · ChartMasterPro Jul. 25, 2017 7 Comments. It Was A Great May For Chip Stocks · SA News Team Jun. 2, 2017 44 Comments. Microchip Tech: Not Much Holding Up ...
15. *MCHP Stock | MICROCHIP TECHNOLOGY Stock Price Today | Markets Insider - Business Insider*
    MCHP : Get the latest Microchip Technology stock price and detailed information including MCHP news, historical charts and realtime prices.
16. *Microchip Technology Stock Price, News & Analysis (NASDAQ:MCHP) - MarketBeat*
    Are you considering Microchip Technology (NASDAQ: MCHP) stock for your investment portfolio? View MCHP's latest stock price, price target, analyst ratings, earnings information, dividend information, headlines ...
17. *Microchip Technology Inc. (MCHP) Stock Updates and News - Fortune*
    Microchip Technology Inc. mchp . NasdaqJanuary 12, 4:00 PM ET. Quotes delayed by at least 15 minutes. Market data provided by ICE Data Services. ICE Limitations. Powered and implemented by FactSet.
18. *MCHP Microchip Technology Incorporated Stock Quote - Finviz*
    Microchip Technology Incorporated develops, manufactures, and sells semiconductor products for various embedded control applications. The company offers general purpose and specialized 8-bit, 16-bit, and 32-bit ...
19. *Buy or Sell MCHP (Microchip Technology Incorporated) stocks?*
    Technical stock analysis for Microchip Technology Incorporated: Several short-term signals are positive and we conclude that the current level may hold a buying opportunity, as there is a fair chance for this stock t...
20. *MCHP:NASDAQ GS Stock Quote - Microchip Technology Inc - Bloomberg Markets*
    Stock analysis for Microchip Technology Inc (MCHP :NASDAQ GS) including stock price, stock chart, company news, key statistics, fundamentals and company profile.
21. *Microchip Technology Inc (MCHP) Stock Analysis - GuruFocus.com*
    Microchip Technology Inc (MCHP) Stock Analysis, 10 Years Financial Analysis, Interactive Charts, Stock Valuations, Value Investing.
22. *MCHP: Dividend Date & History for Microchip Technology*
    Created with Highstock 2.0.4 Zoom From Mar 2, 2017 To Mar 2, 2018 Stock Price Dividend Yield MCHP Price (split-adjusted) MCHP Yield Stock Split Legend (Click to show / hide lines) Apr '17 May '17 Jun '17 Jul ...
23. *MCHP - Stock Quote and Charts for Microchip Technology*
    Microchip Technology (MCHP) stock quote, charts, historical data, financials. Microchip Technology Inc develops, manufactures and sells semiconductor products used for a range of embedded control applications.
24. *MCHP Stock Quote - Microchip Technology Inc Summary | Investopedia*
    Stay up to date with real time MCHP stock quotes, historical charts and the latest financial news and investing data for Microchip Technology Inc.
25. *Microchip Technology Inc, MCHP:NSQ forecasts - FT.com*
    Microchip Technology Inc (MCHP :NSQ) forecasts: consensus recommendations, research reports, share price forecasts, dividends, and earning history and estimates.
26. *Microchip Technology Stock Split History*
    Microchip Technology (MCHP) has 7 splits in our Microchip Technology stock split history database. The first split for MCHP took place on September 14, 1993. This was a 2 for 1 split, meaning for each share of MCHP ...
27. *MCHP Price Target | Microchip Ratings - TipRanks*
    See MCHP price target based on 23 analysts offering 12 month price targets for Microchip (MCHP) in the last 3 months. MCHP's price target is $110.86 with a high estimate of $130.00 and a low estimate ... Reviewer TipRanks
28. *Microchip Technology (NASDAQ:MCHP) Stock Quotes and News | Benzinga*
    View Microchip Technology Incorporated (NASDAQ: MCHP) stock data and charts on Benzinga.
29. *Microchip Technology (MCHP) Stock Message Board - InvestorsHub*
    Company Name: Microchip Technology, Stock Symbol: MCHP , Industry: Computers - Semiconductors, Total Posts: 11, Last Post: 9/30/2017 9:34:55 AM.
30. *MCHP - Google*
    No information is available for this page. Learn why

Related Searches

1. mchp news
2. mchp earnings
3. microchip technology news
4. mu stock
5. microchip technology subsidiaries
6. on stock
7. amat stock
8. adi stock

Figure 3.1: Sample text corpus of Google search results for a stock ticker (Mobile)

This figure presents the text corpus of Google mobile search results for the term "MCHP stock" (i.e. Microchip Technology stock). The search was performed on March 6th, 2018 in an android phone environment. The sentiment scores calculated by VADER are 2.4% negative, 3.6% positive, and 94% neutral.

News

1. What Microchip Technology's $10.15 billion Microsemi deal means for the company, semiconductor industry
2. Microchip Purchase Of Microsemi Gets Thumbs Up From Wall Street | Stock News & Stock Market Analysis
3. Microchip to buy Microsemi for about $8.35 billion

Results

1. *MCHP Stock Price - Microchip Technology Inc. Stock Quote (U.S. ...*
   Microchip Technology Inc. stock price, stock quotes and financial overviews from MarketWatch.
2. *MCHP : Summary for Microchip Technology Incorporat - Yahoo Finance*
   Microchip Technology Inc. (NASDAQ: MCHP) has introduced two new microcontroller families designed with customer innovation in mind. ... It is also projected that this deal could boost MCHP's earnings by 20%-30%.
3. *MCHP 94.47 3.18 3.48% : Microchip Technology Incorporat - Yahoo ...*
   Microchip Technology Inc. (NASDAQ: MCHP) has introduced two new microcontroller families designed with customer innovation in mind. ... Yahoo Finance's Seana Smith and Jared Blikre discuss Bank of America's call for another $6 trillion to be wiped out of equities.
4. *Microchip Technology - MCHP - Stock Price & News | The Motley Fool*
   Real time Microchip Technology (MCHP) stock price quote, stock graph, news & analysis.
5. *Microchip Technology Inc. - NASDAQ:MCHP - Stock Quote & News ...*
   View detailed financial information, real-time news, videos, quotes and analysis on Microchip Technology Inc. (NASDAQ: MCHP). Explore commentary on Microchip Technology Inc. and hear what the experts at TheStreet are saying about MCHP .
6. *MCHP Stock Quote - Microchip Technology Incorporated Common ...*
   Stock quote for Microchip Technology Incorporated Common Stock Common Stock (MCHP) with real-time last sale and extended hours stock prices, company news, charts, and research at Nasdaq.
7. *Microchip Technology Incorporated - MCHP - Stock Price Today - Zacks*
   View Microchip Technology Incorporated MCHP investment & stock information. Get the latest Microchip Technology Incorporated MCHP detailed stock quotes, stock data, Real-Time ECN, charts, stats and more.
8. *MCHP Stock Price & News - Microchip Technology Inc. - Wall Street ...*
   Microchip Technology Inc. Stock - MCHP news, historical stock charts, analyst ratings, financials, and today's Microchip Technology Inc. stock price.
9. *MCHP Microchip Technology Inc. stock and investing information on ...*
   $ MCHP analysts on Estimize are expecting 9.99% YoY revenue growth for Q4, down from 12.83% in Q3 [Reporting 05/09 AMC] http://www.estimize.com/intro/ mchp ?chart=historical&metric_name=revenue&utm_content= MCHP&utm_medium=revenue_growth&utm_source=stocktwits · bizscienceMar 3rd ...
10. *MCHP - Microchip Technology Inc Stock quote - CNNMoney.com*
    Mar 04 / TheStreet.com - Paid Partner Content. Why Microchip Is Spending More Than $10 Billion to Buy Microsemi. Mar 02 / TheStreet.com - Paid Partner Content. Stocks Fall in Premarket Trade. Mar 02 / GuruFocus News - Paid Partner Content. Dow Tumbles Amid Fears of Global Trade War; S&P 500 and Nasdaq Finish ...
11. *MCHP 94.47 +3.48 (Microchip Technology Inc) - Morningstar*
    Morningstar real-time stock quotes, graphs, and independent analysis for MCHP keep you informed. Stay up to date with Microchip Technology Inc stock price and analyst ratings on Morningstar.com.
12. *Microchip Tech Stock Quote: MCHP Stock News, Quotes, Analysis ...*
    MCHP , Microchip Tech - Stock quote performance, technical chart analysis, SmartSelect Ratings, Group Leaders and the latest company headlines.
13. *MCHP Stock Price & News - Microchip Technology Inc. - Barron's*
    View the latest MCHP stock price with Barron's. Including historical share prices, analysis, earnings, cash flow and market valuation for Microchip Technology Inc.
14. *MCHP: 94.47 0.00 (0.00%) - Stock Quote and News - CNBC*
    Get Microchip Technology Inc (MCHP :NASDAQ) real-time stock quotes, news and financial information from CNBC.
15. *MCHP Microchip Technology Incorporated Stock Quote - Finviz*
    Microchip Technology Incorporated develops, manufactures, and sells semiconductor products for various embedded control applications. The company offers general purpose and specialized 8-bit, 16-bit, and 32-bit microcontrollers; and microcontrollers for automotive networking, computing, lighting, power supplies, ...
16. *MCHP:NASDAQ GS Stock Quote - Microchip Technology Inc ...*
    Stock analysis for Microchip Technology Inc (MCHP :NASDAQ GS) including stock price, stock chart, company news, key statistics, fundamentals and company profile.
17. *Microchip Technology Stock Price, News & Analysis (NASDAQ:MCHP)*
    Are you considering Microchip Technology (NASDAQ: MCHP) stock for your investment portfolio? View MCHP's latest stock price, price target, analyst ratings, earnings information, dividend information, headlines, insider trades and SEC filings in real-time at MarketBeat.
18. *MCHP Analysis & News - Microchip Technology Incorporated ...*
    PRO Pick Michael Boyd Nov. 23, 2017 3 Comments. Microchip Technologies: High Risk, Low Reward · ChartMasterPro Jul. 25, 2017 7 Comments. It Was A Great May For Chip Stocks · SA News Team Jun. 2, 2017 44 Comments. Microchip Tech: Not Much Holding Up The Share Price · ChartMasterPro Mar. 23, 2017 16 ...
19. *Microchip Technology Inc. (MCHP) Stock Updates and News - Fortune*
    S&P Index data is the property of Chicago Mercantile Exchange Inc. and its licensors. All rights reserved. 93.45; +1.34 (+1.45 %); n/a Year-to-Date. 1d 1 Day; 5d 5 Day; 1m 1 Month; 6m 6 Month; 1y 1 Year; 5y 5 Year; 10y 10 Year. Reset Chart Chart Options. MCHP . 1. Mar 2. Mar 5. Mar 12:00pm 86 88 90 92 94 96 1. Mar 5.
20. *Buy or Sell MCHP (Microchip Technology Incorporated) stocks?*
    Technical stock analysis for Microchip Technology Incorporated: Several short- term signals are positive and we conclude that the current level may hold a buying opportunity, as there is a fair chance for this stock t...
21. *MCHP - Stock Quote and Charts for Microchip Technology*
    Microchip Technology (MCHP) stock quote, charts, historical data, financials. Microchip Technology Inc develops, manufactures and sells semiconductor products used for a range of embedded control applications. It also provides license Flash-IP solutions.
22. *MCHP Stock | MICROCHIP TECHNOLOGY Stock Price Today ...*
    MCHP : Get the latest Microchip Technology stock price and detailed information including MCHP news, historical charts and realtime prices.
23. *Microchip Technology (NASDAQ:MCHP) Stock Quotes and News ...*
    View Microchip Technology Incorporated (NASDAQ: MCHP) stock data and charts on Benzinga.
24. *Microchip Technology (MCHP) Stock Message Board - InvestorsHub*
    Company Name: Microchip Technology, Stock Symbol: MCHP , Industry: Computers - Semiconductors, Total Posts: 11, Last Post: 9/30/2017 9:34:55 AM.
25. *Microchip Technology Inc (MCHP) Stock Analysis - GuruFocus.com*
    Microchip Technology Inc (MCHP) Stock Analysis, 10 Years Financial Analysis, Interactive Charts, Stock Valuations, Value Investing.
26. *Microchip Technology (MCHP) Stocks Price Quote - Barchart.com*
    Free stocks price quote for MCHP . Latest real-time Bats price quote, charts, financials, technicals and opinions.
27. *MCHP: Dividend Date & History for Microchip Technology*
    Created with Highstock 2.0.4 Zoom From Mar 2, 2017 To Mar 2, 2018 Stock Price Dividend Yield MCHP Price (split-adjusted) MCHP Yield Stock Split Legend (Click to show / hide lines) Apr '17 May '17 Jun '17 Jul '17 Aug '17 Sep '17 Oct '17 Nov '17 Dec '17 Jan '18 Feb '18 Mar '18 1995 2000 2005 2010 2015 $70 $80 $90 ...
28. *Microchip Technology Inc, MCHP:NSQ forecasts - FT.com*
    Microchip Technology Inc (MCHP :NSQ) forecasts: consensus recommendations, research reports, share price forecasts, dividends, and earning history and estimates.
29. *MCHP Stock Quote - Microchip Technology Inc Summary | Investopedia*
    Stay up to date with real time MCHP stock quotes, historical charts and the latest financial news and investing data for Microchip Technology Inc.

Related Searches

1. mchp news
2. mchp earnings
3. mu stock
4. on stock
5. mchp stock news
6. fb stock

Figure 3.2: Sample text corpus of Google search results for a stock ticker (PC)

This figure presents the text corpus of Google PC search results for the same term at the same time as Figure 3.1. The sentiment scores calculated by VADER are 1.6% negative, 4.0% positive, and 94.4% neutral. Notice there are 29 results instead of 30 as Google may sometimes show fewer actual results depending on the page layout.

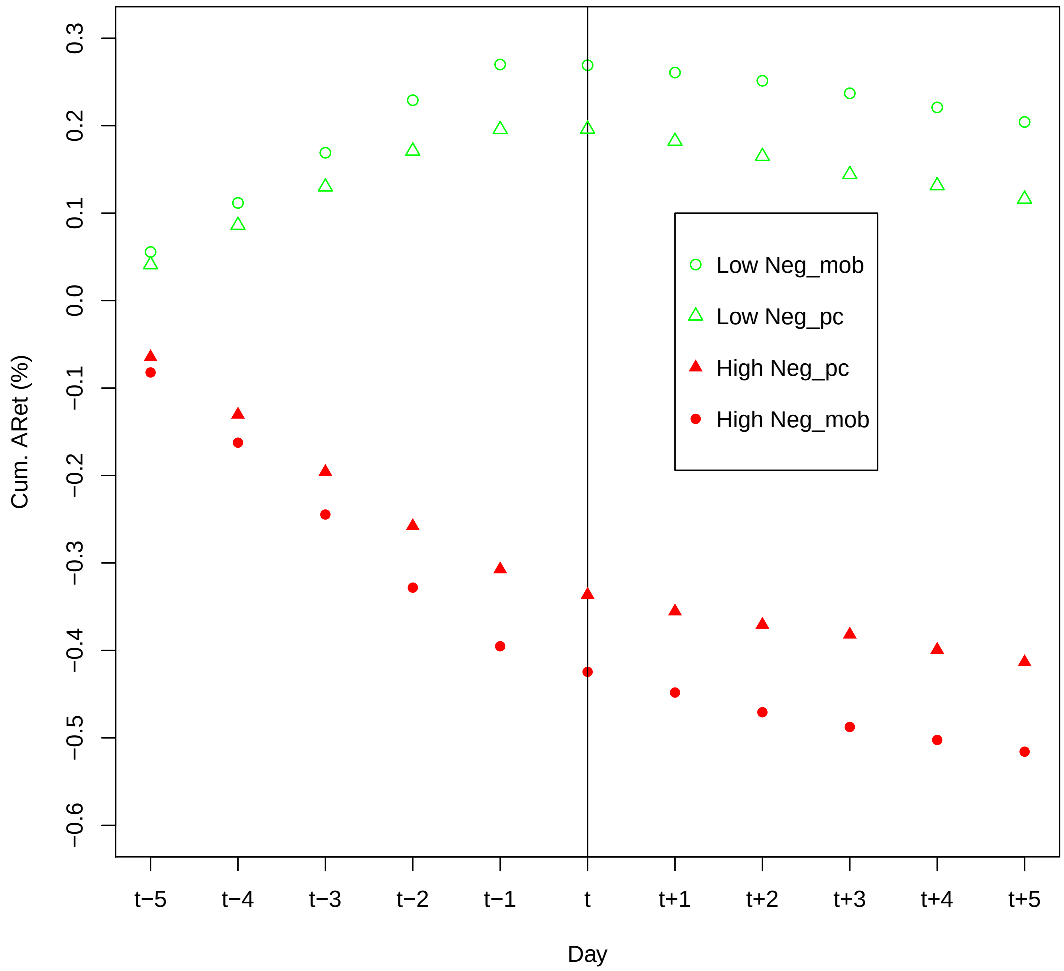

Figure 3.3: Cumulative abnormal returns based on sentiment

This figure shows cumulative abnormal returns of low and high negative sentiment stocks. On trading day t, a stock is in a low (high) negative sentiment subsample if its negative sentiment that day is below (above) the median of negative sentiment scores, which are measured by either mobile or PC results.

3.7 Tables

Table 3.1.A: Summary Statistics for Sentiment Scores (stock/trading day level)

This table reports the summary statistics of sentiment scores evaluated based on the text of Google ticker search results. The *Negative* (*Positive*) score is the total magnitude of negative (positive) intensity divided by the total magnitude of intensity. The negative and positive scores are computed by VADER python library. The sentiment analysis is conducted separately on mobile and PC search results. That is each stock has its mobile-based sentiment scores and its PC-based sentiment scores each day. The sample covers from August 2017 to December 2018 (330 trading days) including 379 U.S.-based common stocks in the S&P 500 index. All numbers are given in percentages.

	Mean	Median	Pctl25	Pctl75	SD
Mobile					
Negative	1.83	1.50	1.00	2.13	1.72
Positive	6.20	5.10	4.10	6.43	4.03
PC					
Negative	1.79	1.40	1.00	2.10	1.85
Positive	6.15	5.00	4.00	6.40	4.17

Table 3.1.B: Summary Statistics for Stock Variables (stock/trading day level)

This table reports the summary statistics of the main variables used in this paper for stock market activity. Abnormal return is the raw return of a firm minus the return of a value-weighted portfolio with similar size, book-to-market ratio, and past returns. Dollar effective spread is computed using constructed NBBO quotes and trades from DTAQ. Turnover is shares traded divided by shares outstanding.

	N	Mean	Median	Pctl25	Pctl75	SD
Abnormal return (%)	116,254	-0.02	0.00	-0.66	0.65	1.42
Dollar effective spread (cents)	116,254	3.53	1.79	1.05	3.98	5.03
Turnover (%)	116,254	0.90	0.68	0.46	1.04	0.92

Table 3.2.A: Predicting Abnormal Returns Using Negative Sentiment

This table reports OLS estimates of the coefficients in Equation 3.1. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. Each sentiment coefficient measures the impact of a 1% increase in negative sentiment on abnormal returns in percentages. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(1.a)	(1.b)	(1.c)
$neg_{mob_{t-1}}$	-0.002		-0.003
	(0.007)		(0.008)
$neg_{mob_{t-2}}$	0.004		0.001
	(0.011)		(0.010)
$neg_{mob_{t-3}}$	0.021**		0.020**
	(0.009)		(0.010)
$neg_{mob_{t-4}}$	-0.003		-0.008
	(0.008)		(0.009)
$neg_{mob_{t-5}}$	0.010		0.010
	(0.008)		(0.009)
$neg_{pc_{t-1}}$		-0.001	-0.001
		(0.008)	(0.010)
$neg_{pc_{t-2}}$		0.008	0.006
		(0.011)	(0.010)
$neg_{pc_{t-3}}$		0.009	-0.0004
		(0.008)	(0.009)
$neg_{pc_{t-4}}$		0.010	0.011
		(0.009)	(0.010)
$neg_{pc_{t-5}}$		0.004	-0.002
		(0.009)	(0.010)
$ARet_{t-1}$	-0.013	-0.013	-0.013
	(0.009)	(0.009)	(0.009)
$ARet_{t-2}$	-0.005	-0.004	-0.005
	(0.008)	(0.008)	(0.008)
$ARet_{t-3}$	0.013	0.013	0.013
	(0.009)	(0.009)	(0.009)
$ARet_{t-4}$	-0.006	-0.006	-0.006
	(0.007)	(0.007)	(0.007)
$ARet_{t-5}$	-0.014*	-0.014*	-0.014*
	(0.007)	(0.007)	(0.007)
Firm FE	Y	Y	Y
Day-of-the-week FE	Y	Y	Y
Observations	112,033	112,033	112,033
R^2	0.004	0.004	0.004
Adjusted R^2	0.001	0.001	0.001
p-value for $\chi^2(5)[Joint_{mob}]$	0.000		0.067
p-value for $\chi^2(5)[Joint_{pc}]$		0.011	0.713
p-value for $\chi^2(10)[Joint_{mob \text{ and } pc}]$			0.007
Sum of five mob lags: $\sum_i \boldsymbol{\beta}_{mob,i}$	0.031***		0.021**
p-value for $\chi^2(1)[\sum_i \boldsymbol{\beta}_{mob,i}]$	0.000		0.046
Sum of five pc lags: $\sum_i \boldsymbol{\beta}_{pc,i}$		0.030***	0.014
p-value for $\chi^2(1)[\sum_i \boldsymbol{\beta}_{pc,i}]$		0.000	0.191

Table 3.2.B: Predicting Abnormal Returns Using Negative Sentiment

This table reports OLS estimates of the coefficients in Equation 3.1 for subsamples sorted based on institutional ownership and size, respectively. Low-Inst (High-Inst) column shows OLS estimates for the quarterly sorted subsample of stocks with less (higher) than median institutional ownership in the $t-1$ quarter. The subsamples cover August 2017 through June 2018. Small-Size (Large-Size) column shows OLS estimates for the daily sorted subsample of stocks with less (higher) than median market cap in the $t-1$ day. The subsamples cover August 2017 through December 2018. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	Low-Inst	High-Inst	Small-Size	Large-Size
$neg_{mob_{t-1}}$	-0.003	-0.024	-0.019	0.009
	(0.011)	(0.019)	(0.015)	(0.009)
$neg_{mob_{t-2}}$	0.006	0.017	0.018	-0.013
	(0.013)	(0.027)	(0.019)	(0.010)
$neg_{mob_{t-3}}$	0.022^{*}	0.028	0.032^{**}	0.010
	(0.012)	(0.019)	(0.016)	(0.012)
$neg_{mob_{t-4}}$	-0.025^{**}	-0.009	-0.015	-0.005
	(0.012)	(0.022)	(0.015)	(0.011)
$neg_{mob_{t-5}}$	0.030^{***}	0.018	0.009	0.008
	(0.011)	(0.018)	(0.017)	(0.011)
$neg_{pc_{t-1}}$	0.004	0.056^{***}	0.007	-0.005
	(0.013)	(0.021)	(0.019)	(0.008)
$neg_{pc_{t-2}}$	0.004	-0.042	-0.004	0.012
	(0.011)	(0.026)	(0.019)	(0.009)
$neg_{pc_{t-3}}$	-0.011	-0.001	-0.021	0.010
	(0.013)	(0.026)	(0.017)	(0.010)
$neg_{pc_{t-4}}$	0.008	0.039^{**}	0.028^{*}	-0.001
	(0.013)	(0.019)	(0.016)	(0.011)
$neg_{pc_{t-5}}$	0.002	-0.021	-0.007	0.004
	(0.013)	(0.020)	(0.019)	(0.011)
$L5(ARet)$	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
Day-of-the-week FE	Y	Y	Y	Y
Observations	41,131	30,408	54,788	60,226
R^2	0.007	0.007	0.005	0.006
Adjusted R^2	0.001	0.001	0.001	0.002
p-value for $\chi^2(5)[Joint_{mob}]$	0.003	0.235	0.049	0.643
p-value for $\chi^2(5)[Joint_{pc}]$	0.948	0.054	0.504	0.495
p-value for $\chi^2(10)[Joint_{mob\text{ and }pc}]$	0.001	0.000	0.082	0.074
Sum of five mob lags: $\sum_i \boldsymbol{\beta}_{mob,i}$	0.031^{**}	0.030	0.025	0.009
p-value for $\chi^2(1)[\sum_i \boldsymbol{\beta}_{mob,i}]$	0.021	0.206	0.125	0.424
Sum of five pc lags: $\sum_i \boldsymbol{\beta}_{pc,i}$	0.006	0.031	0.004	0.020
p-value for $\chi^2(1)[\sum_i \boldsymbol{\beta}_{pc,i}]$	0.707	0.175	0.822	0.110

Table 3.3: Predicting Negative Sentiment Using Abnormal Returns

This table reports OLS estimates of the coefficients in Equation 3.2. The dependent variable is neg_{mob_t} and neg_{pc_t} for models (a) and (b), respectively. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. Each $ARet$ coefficient measures the impact of a 1% increase in abnormal returns on negative sentiment in percentages. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(2.a)	(2.b)
$ARet_{t-1}$	-0.027^{***}	-0.021^{***}
	(0.002)	(0.002)
$ARet_{t-2}$	-0.022^{***}	-0.015^{***}
	(0.002)	(0.001)
$ARet_{t-3}$	-0.015^{***}	-0.010^{***}
	(0.002)	(0.002)
$ARet_{t-4}$	-0.010^{***}	-0.007^{***}
	(0.001)	(0.001)
$ARet_{t-5}$	-0.005^{***}	-0.005^{***}
	(0.001)	(0.001)
$neg_{mob_{t-1}}$	0.413^{***}	0.101^{***}
	(0.009)	(0.007)
$neg_{mob_{t-2}}$	0.132^{***}	-0.008
	(0.007)	(0.006)
$neg_{mob_{t-3}}$	0.070^{***}	-0.016^{***}
	(0.006)	(0.006)
$neg_{mob_{t-4}}$	0.059^{***}	-0.007
	(0.005)	(0.006)
$neg_{mob_{t-5}}$	0.075^{***}	0.002
	(0.007)	(0.006)
$neg_{pc_{t-1}}$	0.104^{***}	0.415^{***}
	(0.009)	(0.011)
$neg_{pc_{t-2}}$	-0.007	0.136^{***}
	(0.007)	(0.009)
$neg_{pc_{t-3}}$	-0.016^{***}	0.064^{***}
	(0.005)	(0.008)
$neg_{pc_{t-4}}$	-0.001	0.061^{***}
	(0.007)	(0.007)
$neg_{pc_{t-5}}$	0.005	0.081^{***}
	(0.007)	(0.006)
Firm FE	Y	Y
Day-of-the-week FE	Y	Y
Observations	112,033	112,033
R^2	0.894	0.911
Adjusted R^2	0.894	0.911
p-value for $\chi^2(5)[Joint_{ARet}]$	0.000	0.000
Sum of five $ARet$ lags: $\sum_i \gamma_i$	$\boxed{-0.079^{***}}$	$\boxed{-0.058^{***}}$
p-value for $\chi^2(1)[\sum_i \gamma_i]$	$\boxed{0.000}$	$\boxed{0.000}$

Table 3.4: Predicting Effective Spreads Using Negative Sentiment

This table reports OLS estimates of the coefficients in equation 3.3. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. The prefix AD denotes absolute difference between current sentiment and mean of daily sentiments over the past trading week. On a given platform, each $ADneg$ coefficient measures the impact of a 1% increase in absolute difference of negative sentiment on effective spreads in percentages. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(3.a)	(3.b)	(3.c)
$neg_{mob_{t-1}}$	-0.316^{**}		-0.367^{***}
	(0.126)		(0.116)
$neg_{mob_{t-2}}$	-0.002		0.069
	(0.106)		(0.109)
$neg_{mob_{t-3}}$	-0.104		-0.077
	(0.106)		(0.105)
$neg_{mob_{t-4}}$	-0.140		-0.077
	(0.106)		(0.113)
$neg_{mob_{t-5}}$	0.002		-0.034
	(0.110)		(0.115)
$ADneg_{mob_{t-1}}$	-0.381^{**}		-0.301^{*}
	(0.168)		(0.174)
$ADneg_{mob_{t-2}}$	0.098		0.238
	(0.149)		(0.154)
$ADneg_{mob_{t-3}}$	-0.417^{***}		-0.329^{**}
	(0.158)		(0.155)
$ADneg_{mob_{t-4}}$	-0.259^{*}		-0.243^{*}
	(0.146)		(0.143)
$ADneg_{mob_{t-5}}$	-0.245		-0.171
	(0.155)		(0.150)
$neg_{pc_{t-1}}$		-0.077	0.138
		(0.133)	(0.121)
$neg_{pc_{t-2}}$		-0.144	-0.137
		(0.109)	(0.109)
$neg_{pc_{t-3}}$		-0.072	-0.006
		(0.118)	(0.119)
$neg_{pc_{t-4}}$		-0.214^{**}	-0.153
		(0.107)	(0.115)
$neg_{pc_{t-5}}$		0.065	0.114
		(0.103)	(0.106)
$ADneg_{pc_{t-1}}$		-0.297^{*}	-0.180
		(0.169)	(0.176)
$ADneg_{pc_{t-2}}$		-0.280^{*}	-0.320^{**}
		(0.153)	(0.153)
$ADneg_{pc_{t-3}}$		-0.336^{**}	-0.210
		(0.144)	(0.137)
$ADneg_{pc_{t-4}}$		-0.049	0.077
		(0.138)	(0.133)
$ADneg_{pc_{t-5}}$		-0.294^{**}	-0.209^{*}
		(0.136)	(0.124)
$\Delta ESpd_{t-1}$	-0.562^{***}	-0.562^{***}	-0.562^{***}
	(0.032)	(0.032)	(0.032)
$\Delta ESpd_{t-2}$	-0.371^{***}	-0.371^{***}	-0.371^{***}
	(0.027)	(0.027)	(0.027)
$\Delta ESpd_{t-3}$	-0.227^{***}	-0.227^{***}	-0.227^{***}
	(0.025)	(0.025)	(0.025)
$\Delta ESpd_{t-4}$	-0.144^{***}	-0.143^{***}	-0.144^{***}
	(0.026)	(0.026)	(0.026)
$\Delta ESpd_{t-5}$	-0.070^{***}	-0.070^{***}	-0.070^{***}
	(0.023)	(0.023)	(0.023)
Firm FE	Y	Y	Y
Day-of-the-week FE	Y	Y	Y
Observations	97,123	97,123	97,123
R^2	0.247	0.247	0.247
Adjusted R^2	0.244	0.244	0.244
p-value for $\chi^2(5)[Joint_{neg_{mob}}]$	0.034		0.015
p-value for $\chi^2(5)[Joint_{neg_{pc}}]$		0.079	0.334
p-value for $\chi^2(5)[Joint_{ADneg_{mob}}]$	0.002		0.012
p-value for $\chi^2(5)[Joint_{ADneg_{pc}}]$		0.000	0.021

Table 3.5: Predicting Negative Sentiment Using Effective Spreads

This table reports OLS estimates of the coefficients in equation 3.4. The dependent variable is neg_{mob_t} and neg_{pc_t} for models (a) and (b), respectively. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. Each $\Delta ESpd$ coefficient measures the impact of a 1% increase in dollar effective spreads on negative sentiment in percentages. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(4.a)	(4.b)
$\Delta ESpd_{t-1}$	0.001^{**}	0.001^{***}
	(0.0002)	(0.0002)
$\Delta ESpd_{t-2}$	0.001^{***}	0.001^{***}
	(0.0003)	(0.0002)
$\Delta ESpd_{t-3}$	0.001^{***}	0.001^{***}
	(0.0002)	(0.0002)
$\Delta ESpd_{t-4}$	0.001^{***}	0.001^{***}
	(0.0002)	(0.0002)
$\Delta ESpd_{t-5}$	0.0005^{***}	0.0004^{***}
	(0.0002)	(0.0001)
$neg_{mob_{t-1}}$	0.421^{***}	0.108^{***}
	(0.010)	(0.008)
$neg_{mob_{t-2}}$	0.132^{***}	-0.011^{*}
	(0.008)	(0.007)
$neg_{mob_{t-3}}$	0.072^{***}	-0.017^{**}
	(0.006)	(0.007)
$neg_{mob_{t-4}}$	0.059^{***}	-0.006
	(0.006)	(0.006)
$neg_{mob_{t-5}}$	0.070^{***}	-0.002
	(0.007)	(0.006)
$neg_{pc_{t-1}}$	0.105^{***}	0.418^{***}
	(0.010)	(0.012)
$neg_{pc_{t-2}}$	-0.005	0.136^{***}
	(0.007)	(0.010)
$neg_{pc_{t-3}}$	-0.020^{***}	0.063^{***}
	(0.006)	(0.007)
$neg_{pc_{t-4}}$	-0.002	0.065^{***}
	(0.007)	(0.007)
$neg_{pc_{t-5}}$	0.004	0.082^{***}
	(0.007)	(0.005)
Firm FE	Y	Y
Day-of-the-week FE	Y	Y
Observations	98,637	98,637
R^2	0.893	0.912
Adjusted R^2	0.893	0.911
p-value for $\chi^2(5)[Joint_{\Delta ESpd}]$	0.000	0.000
Sum of five $\Delta ESpd$ lags: $\sum_i \gamma_i$	$\boxed{0.004^{***}}$	$\boxed{0.004^{***}}$
p-value for $\chi^2(1)[\sum_i \gamma_i]$	$\boxed{0.000}$	$\boxed{0.000}$

Table 3.6: Predicting Volume Using Negative Sentiment

This table reports OLS estimates of the coefficients in equation 3.5. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. The prefix AD denotes absolute difference between current sentiment and mean of daily sentiments over the past trading week. On a given platform, each $ADneg$ coefficient measures the impact of a 1% increase in absolute difference of negative sentiment on detrended log turnover. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(5.a)	(5.b)	(5.c)
$neg_{mob_{t-1}}$	-0.002		-0.001
	(0.003)		(0.002)
$neg_{mob_{t-2}}$	-0.002		-0.003
	(0.002)		(0.002)
$neg_{mob_{t-3}}$	0.001		0.001
	(0.002)		(0.002)
$neg_{mob_{t-4}}$	-0.005^{*}		-0.004
	(0.002)		(0.003)
$neg_{mob_{t-5}}$	-0.002		-0.001
	(0.002)		(0.003)
$ADneg_{mob_{t-1}}$	0.006		0.006^{*}
	(0.004)		(0.004)
$ADneg_{mob_{t-2}}$	-0.008^{**}		-0.006
	(0.004)		(0.004)
$ADneg_{mob_{t-3}}$	-0.004		-0.005
	(0.004)		(0.003)
$ADneg_{mob_{t-4}}$	-0.002		-0.003
	(0.004)		(0.004)
$ADneg_{mob_{t-5}}$	-0.004		-0.002
	(0.004)		(0.004)
$neg_{pc_{t-1}}$		-0.003	-0.002
		(0.003)	(0.003)
$neg_{pc_{t-2}}$		0.001	0.003
		(0.002)	(0.002)
$neg_{pc_{t-3}}$		0.0002	0.0003
		(0.002)	(0.003)
$neg_{pc_{t-4}}$		-0.004^{*}	-0.002
		(0.002)	(0.003)
$neg_{pc_{t-5}}$		-0.002	-0.001
		(0.002)	(0.003)
$ADneg_{pc_{t-1}}$		0.003	0.002
		(0.004)	(0.004)
$ADneg_{pc_{t-2}}$		-0.009^{**}	-0.007^{*}
		(0.004)	(0.004)
$ADneg_{pc_{t-3}}$		-0.001	0.001
		(0.004)	(0.003)
$ADneg_{pc_{t-4}}$		0.002	0.004
		(0.004)	(0.004)
$ADneg_{pc_{t-5}}$		-0.009^{**}	-0.008^{**}
		(0.004)	(0.003)
Vlm_{t-1}	0.440^{***}	0.440^{***}	0.440^{***}
	(0.027)	(0.027)	(0.027)
Vlm_{t-2}	0.087^{***}	0.087^{***}	0.087^{***}
	(0.016)	(0.016)	(0.016)
Vlm_{t-3}	0.042^{***}	0.042^{***}	0.042^{***}
	(0.016)	(0.016)	(0.016)
Vlm_{t-4}	0.012	0.012	0.012
	(0.019)	(0.019)	(0.019)
Vlm_{t-5}	-0.023	-0.023	-0.023
	(0.016)	(0.016)	(0.016)
Firm FE	Y	Y	Y
Day-of-the-week FE	Y	Y	Y
Observations	97,126	97,126	97,126
R^2	0.265	0.265	0.265
Adjusted R^2	0.262	0.262	0.262
p-value for $\chi^2(5)[Joint_{neg_{mob}}]$	0.091		0.362
p-value for $\chi^2(5)[Joint_{neg_{pc}}]$		0.238	0.817
p-value for $\chi^2(5)[Joint_{ADneg_{mob}}]$	0.060		0.165
p-value for $\chi^2(5)[Joint_{ADneg_{pc}}]$		0.020	0.055

Table 3.7: Predicting Negative Sentiment Using Volume

This table reports OLS estimates of the coefficients in equation 3.6. The dependent variable is neg_{mob_t} and neg_{pc_t} for models (a) and (b), respectively. The negative sentiment scores (neg) are calculated separately using mobile and PC results, denoted by mob and pc subscripts respectively. Each Vlm coefficient describes the impact of a 1 unit increase in detrended log turnover on negative sentiment in percentages. Robust standard errors double-clustered by firm and day are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	(6.a)	(6.b)
Vlm_{t-1}	0.051^{***}	0.044^{***}
	(0.008)	(0.008)
Vlm_{t-2}	0.014^{*}	0.017^{**}
	(0.008)	(0.008)
Vlm_{t-3}	-0.003	-0.004
	(0.007)	(0.007)
Vlm_{t-4}	0.009	0.004
	(0.009)	(0.008)
Vlm_{t-5}	-0.008	-0.011^{*}
	(0.009)	(0.007)
$neg_{mob_{t-1}}$	0.422^{***}	0.107^{***}
	(0.009)	(0.007)
$neg_{mob_{t-2}}$	0.130^{***}	-0.011^{*}
	(0.007)	(0.006)
$neg_{mob_{t-3}}$	0.069^{***}	-0.011^{*}
	(0.007)	(0.006)
$neg_{mob_{t-4}}$	0.060^{***}	-0.011^{*}
	(0.006)	(0.007)
$neg_{mob_{t-5}}$	0.074^{***}	0.001
	(0.007)	(0.006)
$neg_{pc_{t-1}}$	0.113^{***}	0.429^{***}
	(0.010)	(0.011)
$neg_{pc_{t-2}}$	-0.010	0.131^{***}
	(0.008)	(0.011)
$neg_{pc_{t-3}}$	-0.020^{***}	0.056^{***}
	(0.005)	(0.008)
$neg_{pc_{t-4}}$	-0.004	0.066^{***}
	(0.008)	(0.008)
$neg_{pc_{t-5}}$	0.001	0.076^{***}
	(0.007)	(0.006)
Firm FE	Y	Y
Day-of-the-week FE	Y	Y
Observations	97,191	97,191
R^2	0.893	0.911
Adjusted R^2	0.893	0.911
p-value for $\chi^2(5)[Joint_{Vlm}]$	0.000	0.000
Sum of five Vlm lags: $\sum_i \gamma_i$	0.063^{***}	0.050^{***}
p-value for $\chi^2(1)[\sum_i \gamma_i]$	0.000	0.000

APPENDIX A

APPENDIX

A.1 Platform Figures

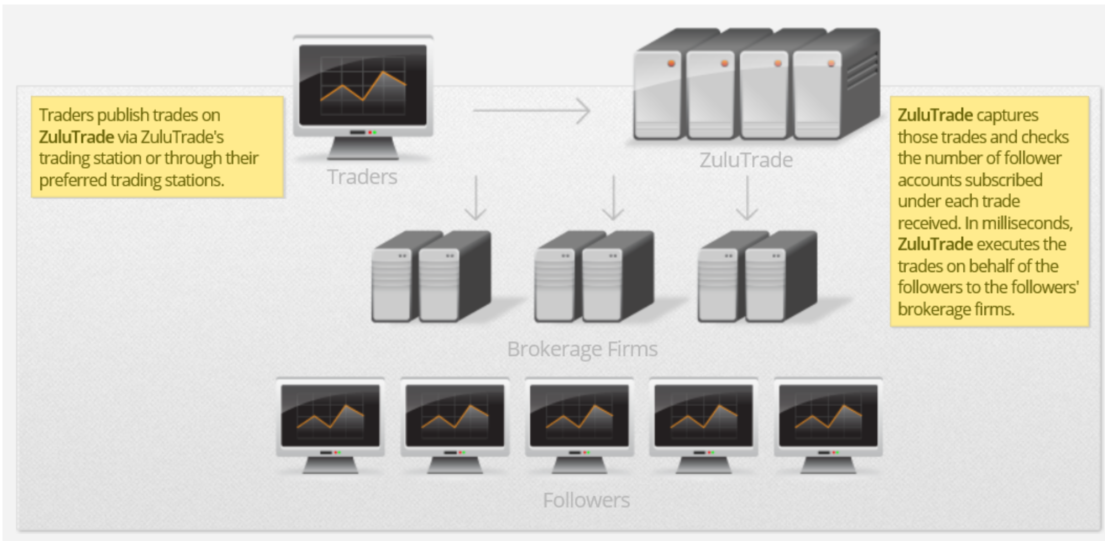

Figure A.1: Trade copying process

This flowchart explains how Zulutrade copies trades into followers' brokerage accounts when traders place the trades. The process is entirely automated. The followers must decide in advance the following settings tailored to each trader (see Figure A.4).

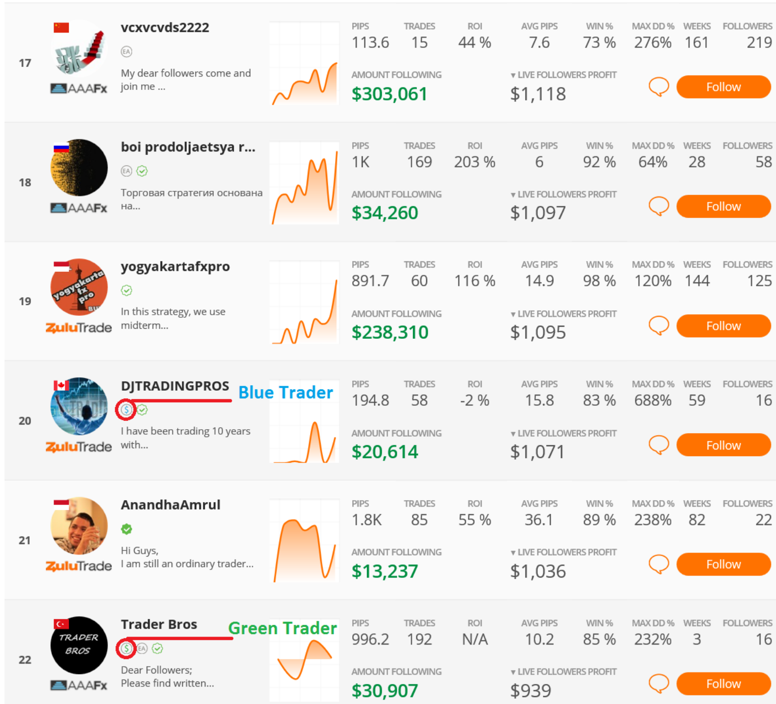

Figure A.2: Trader list and skin-in-the-game (green/blue dollar badge)

This figure illustrates the trader list page of the platform. It provides some summary statistics about the traders and their followers' fund performance and risk. Zulutrade indicates with a dollar-sign $ badge that a trader is putting his/her own real money into his/her own trading. The dollar-sign badge comes in green and blue colors which both imply skin-in-the-game.

A green trader must link a real brokerage account to his/her own Zulutrade trader account. A blue trader is both a trader and a follower of his/her own trading, having both a trader account and a follower account on Zulutrade: a demo/practice (play money) account is linked to the trader account, which is followed by his/her own follower account.

The traders listed here are ordered by last-month profit of followers' fund. Zulutrade also allows a viewer to choose other criteria to sort the traders such as its own proprietary measure ZuluRank.

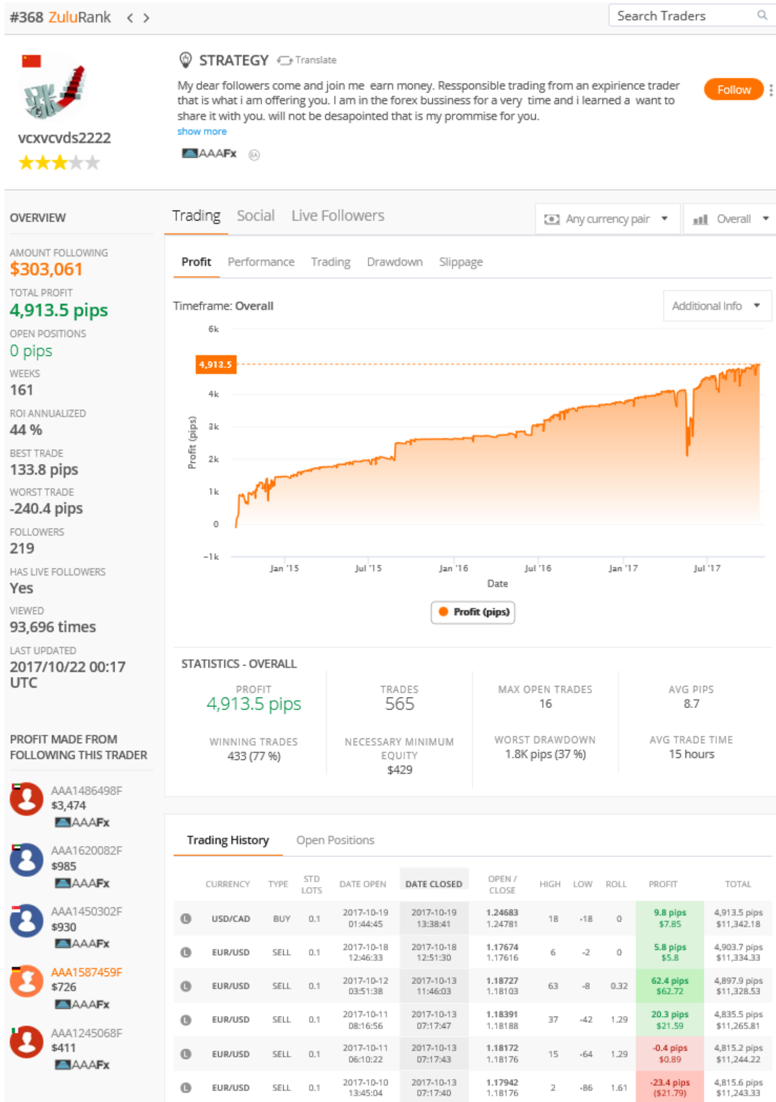

Figure A.3: Trader profile and trading record

This figure illustrates a trader's profile page with trading strategy description, summary statistics and followers' feedbacks. The trading history encompasses the trader's complete time-stamped transactional-level trading record since registered on the platform as well as the current open positions.

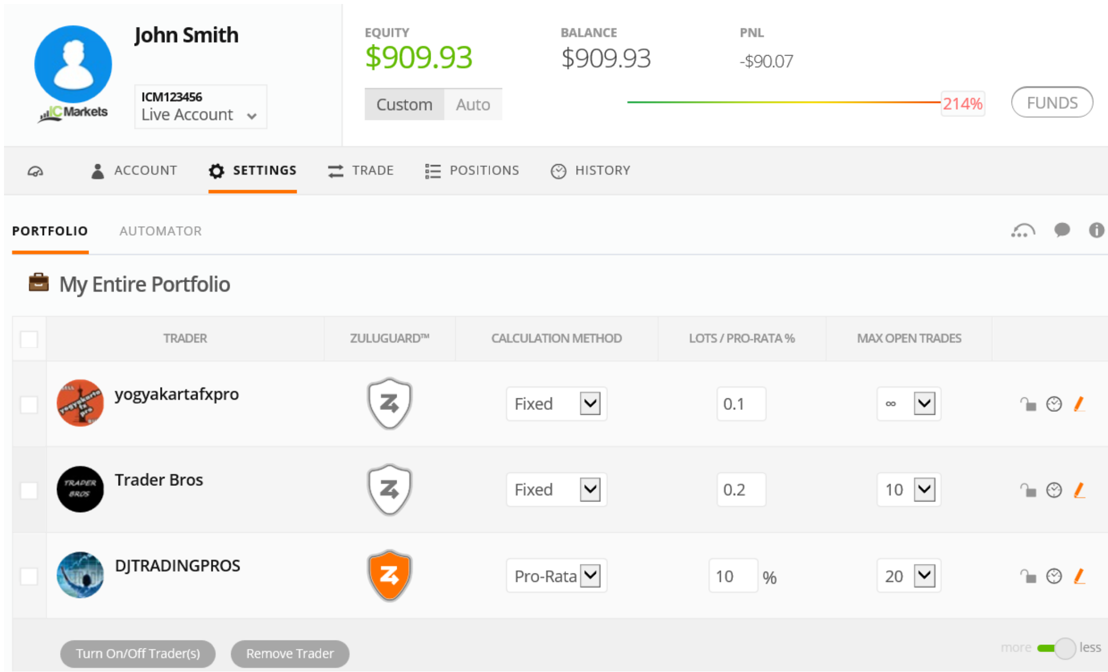

Figure A.4: Follower account dashboard

This figure shows the dashboard user interface for a sample follower account. The follower here links his/her brokerage account from the broker IC Markets with the Zulutrade follower account. The follower sets the size of each trade from traders in advance using a either fixed lot or proportional (Pro-Rata) lot size. For the fixed lot approach, suppose the follower uses 0.1 lot, then whenever a trader makes a new trade whether it is 1 lot, 0.5 lot, or 0.1 lot, the follower will always take 0.1 lot of that trade. For the proportional lot approach, suppose the follower uses 10%, then the follower will take 10% of the lot used by the trader, hence taking 0.1 lot, 0.05 lot or 0.01 lot for trader's 1 lot, 0.5 lot, or 0.1 lot trades respectively. For each trader, the follower can also customize further to set the size of trade by currency pairs (e.g. EUR/USD, GBP/USD).

TRADING HISTORY OPEN POSITIONS Select currency pairs

TRADER	CURRENCY	TYPE	STD LOTS	DATE OPEN	DATE CLOSED	OPEN/ CLOSE	HIGH	LOW	ROLL	PROFIT	TOTAL
yogyakartafxpro	GBP/AUD	BUY	0.01	2017-10-19 18:02:47	2017-10-19 18:50:22	1.67185 1.67248	10	-2	0	6.3 pips €0.24	-98 pips -€17.95
yogyakartafxpro	GBP/USD	SELL	0.01	2017-10-18 16:42:57	2017-10-18 17:11:19	1.31982 1.31903	9	-2	0	7.9 pips €0.49	-104 pips -€18.19
yogyakartafxpro	GBP/USD	SELL	0.01	2017-10-18 16:42:01	2017-10-18 17:11:19	1.31973 1.31903	8	-3	0	7 pips €0.41	-112 pips -€18.68
yogyakartafxpro	EUR/USD	BUY	0.01	2017-10-17 06:31:42	2017-10-17 07:59:56	1.17618 1.17726	11	-7	0	10.8 pips €0.74	-119 pips -€19.09
yogyakartafxpro	EUR/USD	BUY	0.01	2017-10-17 06:30:42	2017-10-17 07:59:56	1.17607 1.1773	12	-6	0	12.3 pips €0.86	-130 pips -€19.83
yogyakartafxpro	EUR/USD	SELL	0.01	2017-10-12 04:59:00	2017-10-12 06:45:02	1.18758 1.18667	10	-4	0	9.1 pips €0.59	-142 pips -€20.69

Figure A.5: Follower trading record

This figure shows the trading history of a follower following the trader yogyakartafxpro. The trading record is at transactional-level and time-stamped; all trades are linked with the user names of the traders who have initiated the trades.

A.2 Trading Instruments

The retail spot forex market is purely speculative in nature. The trading instruments offered to retail investors are fairly complex and opaque. Though most retail brokers claim investors are trading spot forex, there is no physical exchange of currencies ever taking place. This is different from conventional spot contract which requires physical currency delivery generally within two business days. Depending on the broker, there are two types of contracts offered to retail clients: rolling spot forex contracts and contract for difference (CFD) forex contracts.

Rolling spot forex contract is an over-the-counter (OTC) derivative issued by a broker and is traded between the retail investor and the broker. Technically, it is a combination of a spot and a derivative forex transaction: a sequence of spot forex transactions which automatically roll over onto the next day until the investor decides to terminate the contract. It is the main instrument traded in the retail spot forex market. Utilizing leverage, the contract is traded on margin and is non-transferrable, i.e. a forex contract bought from a specific broker cannot be sold to another broker, trader or market maker (Fassas, 2015).

A CFD is a contract in which the buyer receives the difference between the current value of an asset and its value at contract time. If the difference is negative, then the buyer pays the seller the difference instead. CFDs also allow the buyer to hold short positions, which means the buyer realizes loss when the underlying asset increases in value. CFDs also employ leverage with margin requirement, exposing buyers to the potential risks and rewards of holding an investment without actually owning it. Unlike futures contracts, CFDs have no fixed contract size or expiry date. In addition, CFDs can be traded in all kinds of contracts covering equity indices, energy, and metals as well as novel assets like Bitcoin. In the U.S., CFDs are regulated under Title VII of the Dodd Frank Act, which defines CFDs as either a swap or securities based swap. Due to restrictions by the Securities and Exchange Commission (SEC) on OTC financial instruments, trading in the CFD market is not available to retail U.S. residents[1]. However rolling spot forex trading is still allowed currently. This is mainly because there is no central exchange for the forex market, exempting forex from the Title VII regulations that apply to CFD contracts.

Most retail forex brokers provide clients with demo/paper trading accounts with customizable amount of play money. This enables traders to simultaneously test as many strategies as they want in simulated live market conditions without any actual real money involved.

A.3 Platform Business Models & Regulation

A typical social trading platform tracks traders' historical and current trades, and publishes those performance statistics to help followers make informed following decisions. The platform obtains traders' trading records directly from traders' brokers to avoid any manipulation. Once having chosen the traders, the followers receive any new trades published by the traders. These new trades are automatically executed in real-time by the platform into the followers' brokerage accounts. There are two common types of business models.

[1]For this regulatory reason, eToro, a major social trading platform popular in Europe, is forbidden to accept any U.S. clients since it offers CFD trading only

- Subscription based: a trader lists and decides the subscription price (e.g. weekly or monthly) for offering his/her services on the platform. The subscription fees are generally paid in advance and nonrefundable independent of traders' eventual performance or followers' profit or loss. The platform makes money from taking a share of trader's subscription revenue.
- Trading volume based: the platform serves as an introducing broker (IB) to the followers' brokers and receive fees from these brokers whenever the platform executes trades on behalf of the followers. In this case, followers do not see any visible fees paid to the platform since their brokers have already bundled these fees into the trading transaction costs. Then traders obtain a share of platform's revenue for their services. The platform pays more to a trader who generates more gross trading volume that sum over all of the trader's followers. Essentially in this setting, traders are indirectly paid by the followers via the platform, but the eventual payout each trader get (which might be 0) is arbitrarily determined by the platform considering factors like trader's realized performance and overall followers' profitability.

A platform may undertake extensive in-house analysis of traders' performance and attribute data as basis for the ranking of traders on the platform [2]. Despite the growing interest in social trading, the regulatory requirements for the platform or the traders are uncertain. Currently, the traders are unregulated and may not have any professional license. The platform only needs to be regulated if acting as an IB. A platform that just sells subscription service is at a legal

[2] See Lee and Ma (2015) for a discussion on designing ranking algorithm tailored to social trading

gray area (Doering, Neumann, and Paul, 2015).

A.4 Forex Terms and Profit/Loss Calculation

Suppose EUR/USD is trading at 1.1498/1.1500 (Bid/Ask)

EUR is the base currency and USD is the quote currency.

When you buy/long a currency pair, you are buying the base currency while simultaneously selling the quote currency. You want the base currency to rise in value and then you would sell it back at a higher price.

When you sell/short a currency pair, you are selling the base currency while simultaneously buying the quote currency. You want the base currency to fall in value and then you would buy it back at a lower price.

The ask or offer is the best available price at which you will buy a currency pair from the market. Alternatively, it is the price at which your broker will sell the base currency to you in exchange for the quote currency.

The bid is the best available price at which you will sell a currency pair to the market. Alternatively, it is the price at which your broker will buy the base currency from you in exchange for the quote currency.

The bid price is always lower than the ask price because a broker will not sell the base currency (ask price) for lower than the price the broker is willing to pay for it (bid price).

The difference between the bid and the ask prices is known as the spread, which represents the transaction cost of making a trade.

Here, if you think EUR/USD will go up or euro will appreciate against dollar, you will buy euros at 1.1500. If you think EUR/USD will go down or euro will depreciate against dollar, you will sell euros at 1.1498.

Example of Calculating Gains in Forex

As before, EUR/USD is trading at 1.1498/1.1500 (Bid/Ask).

1 standard lot = 100,000 units of base currency
1 pip = 0.0001 movement in exchange rate (but 0.01 for JPY currency pairs)

To buy EUR/USD with 1 lot or 100,000 units, you buy 100,000 euros at 1.1500 with $115,000.

Later EUR/USD moves up to 1.1502/1.1504 (Bid/Ask).

You can exit the position by selling 100,000 euros at 1.1502 and receive $115,020.

Exchange rate movement = 1.1502-1.1500 = 0.0002 = 2 pips

Profit = $115,020-$115,000 = $20

The value of a pip = Profit/Pip gains = $10/pip

Generally the smallest lot size is 1,000 units or 0.01 lot. The value of a pip is always $10/lot or $0.1/0.01 lot when the U.S. dollar is the quote currency of a currency pair (e.g. GBP/USD, AUD/USD, NZD/USD).

BIBLIOGRAPHY

Admati, Anat R. and Paul Pfleiderer. 1988. "A Theory of Intraday Patterns: Volume and Price Variability." *The Review of Financial Studies* 1 (1):3–40.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik. 2009. "Role of Managerial Incentives and Discretion in Hedge Fund Performance." *The Journal of Finance* 64 (5):2221–2256.

Andonov, Aleksandar and Joshua D. Rauh. 2018. "The Return Expectations of Institutional Investors." URL: `https://ssrn.com/abstract=3091976`.

Antweiler, Werner and Murray Z. Frank. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards." *The Journal of Finance* 59 (3):1259–1294.

Barber, Brad and Douglas Loeffler. 1993. "The "Dartboard" Column: Second-Hand Information and Price Pressure." *Journal of Financial and Quantitative Analysis* 28 (02):273–284.

Barber, Brad M. and Terrance Odean. 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics* 116 (1):261–292.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015. "X-CAPM: An extrapolative capital asset pricing model." *Journal of Financial Economics* 115 (1):1–24.

———. 2018. "Extrapolation and bubbles." *Journal of Financial Economics* 129 (2):203–227.

Benartzi, Shlomo. 2001. "Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock." *The Journal of Finance* 56 (5):1747–1764.

Berk, Jonathan B. and Richard Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112 (6):1269–1295.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar. 2009. "How Noise Trading Affects Markets: An Experimental Analysis." *The Review of Financial Studies* 22 (6):2275–2302.

Brown, Stephen, William Goetzmann, Bing Liang, and Christopher Schwarz. 2008. "Mandatory Disclosure and Operational Risk: Evidence from Hedge Fund Registration." *The Journal of Finance* 63 (6):2785–2815.

Campbell, John, Sanford J. Grossman, and Jiang Wang. 1993. "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics* 108 (4):905–939.

Campbell, John Y., Andrew W. Lo, A. Craig MacKinlay, Petr Adamek, and Luis M. Viceira. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Cassella, Stefano and Huseyin Gulen. 2018. "Extrapolation Bias and the Predictability of Stock Returns by Price-Scaled Variables." *The Review of Financial Studies* 31 (11):4345–4397.

Chang, Tom Y., David H. Solomon, and Mark Westerfield. 2016. "Looking for Someone to Blame: Delegation, Cognitive Dissonance, and the Disposition Effect." *Journal of Finance* 71 (1):267–302.

Chen, Hailiang, Prabuddha De, Yu (Jeffrey) Hu, and Byoung-Hyoun Hwang. 2014. "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media." *The Review of Financial Studies* 27 (5):1367–1403.

Chevalier, Judith and Glenn Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." *Journal of Political Economy* 105 (6):1167–1200.

Cremers, Martijn, Joost Driessen, Pascal Maenhout, and David Weinbaum. 2009. "Does Skin in the Game Matter? Director Incentives and Governance in the Mutual Fund Industry." *The Journal of Financial and Quantitative Analysis* 44 (6):1345–1373.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. "In Search of Attention." *Journal of Finance* 66 (5):1461–1499.

Da, Zhi, Umit G. Gurun, and Mitch Warachka. 2014. "Frog in the Pan: Continuous Information and Momentum." *Review of Financial Studies* 27 (7):2171–2218.

Da, Zhi, Xing Huang, and Lawrence Jin. 2018. "Extrapolative Beliefs in the Cross-Section: What Can We Learn from the Crowds?" URL: `https://ssrn.com/abstract=3144849`.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers. 1997. "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks." *The Journal of Finance* 52 (3):1035–1058.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990a. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98 (4):703–738.

———. 1990b. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance* 45 (2):379–395.

Doering, Philipp, Sascha Neumann, and Stephan Paul. 2015. "A Primer on Social Trading Networks – Institutional Aspects and Empirical Evidence."

Dunaway, Johanna, Kathleen Searles, Mingxiao Sui, and Newly Paul. 2018. "News Attention in a Mobile Era." *Journal of Computer-Mediated Communication* 23 (2):107–124.

Engelberg, Joseph E. and Christopher A. Parsons. 2011. "The Causal Impact of Media in Financial Markets." *The Journal of Finance* 66 (1):67–97.

Ertan, Aytekin, Stephen A. Karolyi, Peter W. Kelly, and Robert Stoumbos. 2017. "Individual Investor Overextrapolation." URL: `https://ssrn.com/abstract=2720573`.

Evans, Allison L. 2008. "Portfolio Manager Ownership and Mutual Fund Performance." *Financial Management* 37 (3):513–534.

Fama, Eugene and James D MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3).

Fassas, Athanasios P. 2015. "Forex Retail Online Trading: Why does this business continue to grow?: A Pitch."

García, Diego. 2012. "Sentiment during Recessions." *The Journal of Finance* 68 (3):1267–1300.

Glosten, Lawrence R. and Paul R. Milgrom. 1985. "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders." *Journal of Financial Economics* 14 (1):71–100.

Granger, Clive. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3):424–38.

Greenwood, Robin and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *The Review of Financial Studies* 27 (3):714–746.

Guo, Li, Feng Shi, and Jun Tu. 2016. "Textual analysis and machine leaning: Crack unstructured data in finance and accounting." *The Journal of Finance and Data Science* 2 (3):153–170.

Gupta, Arpit and Kunal Sachdeva. 2018. "Skin or Skim? Inside Investment and Hedge Fund Performance." URL: `https://ssrn.com/abstract=2983030`.

Holden, Craig W. and Stacey Jacobsen. 2014. "Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions." *Journal of Finance* 69 (4):1747–1785.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. 2011. "Risk Shifting and Mutual Fund Performance." *The Review of Financial Studies* 24 (8):2575–2616.

Huberman, Gur and Tomer Regev. 2001. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar." *The Journal of Finance* 56 (1):387–396.

Hutto, C.J. and Eric Gilbert. 2014. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." In *8th international AAAI conference on weblogs and social media (ICWSM)*.

Jin, Lawrence J. and Pengfei Sui. 2018. "Asset Pricing with Return Extrapolation." URL: `https://ssrn.com/abstract=3045658`.

Joseph, Kissan, M. Babajide Wintoki, and Zelin Zhang. 2011. "Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search." *International Journal of Forecasting* 27 (4):1116–1127.

Khorana, Ajay, Henri Servaes, and Lei Wedge. 2007. "Portfolio manager ownership and fund performance." *Journal of Financial Economics* 85 (1):179–204.

Kim, Ki Joon and S. Shyam Sundar. 2016. "Mobile Persuasion: Can Screen Size and Presentation Mode Make a Difference to Trust?" *Human Communication Research* 42 (1):45–70.

Lee, Charles and Mark Ready. 1991. "Inferring Trade Direction from Intraday Data." *Journal of Finance* 46 (2):733–46.

Lee, Charles M.C. and Bhaskaran Swaminathan. 2000. "Price Momentum and Trading Volume." *Journal of Finance* 55 (5):2017–2069.

Lee, Woonyeol and Qiang Ma. 2015. "Whom to Follow on Social Trading Services? A System to Support Discovering Expert Traders."

Liu, Shuming. 2015. "Investor Sentiment and Stock Market Liquidity." *Journal of Behavioral Finance* 16 (1):51–67.

Loughran, Tim and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66 (1):35–65.

———. 2016. "Textual Analysis in Accounting and Finance: A Survey." *Journal of Accounting Research* 54 (4):1187–1230.

Ma, Linlin and Yuehua Tang. 2014. "Portfolio Manager Ownership and Mutual Fund Risk Taking." URL: `https://ssrn.com/abstract=2024119`.

Malmendier, Ulrike and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *The Quarterly Journal of Economics* 126 (1):373–416.

Miller, Edward M. 1977. "Risk, Uncertainty, and Divergence of Opinion." *The Journal of Finance* 32 (4):1151–1168.

Newey, Whitney and Kenneth West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3).

Ozik, Gideon and Ronnie Sadka. 2015. "Skin in the Game versus Skimming the Game: Governance, Share Restrictions, and Insider Flows." *Journal of Financial and Quantitative Analysis* 50 (6):1293–1319.

Ribeiro, Filipe N., Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabrício Benevenuto. 2016. "SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods." *EPJ Data Science* 5 (1).

Sirri, Erik R. and Peter Tufano. 1998. "Costly Search and Mutual Fund Flows." *The Journal of Finance* 53 (5):1589–1622.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance* 62 (3):1139–1168.