THREE ESSAYS IN EMPIRCAL ASSET PRICING

XINYAO ZHOU

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ADMINISTRATION
YORK UNIVERSITY
TOTRONTO, ONTARIO

JUNE, 2022

ABSTRACT

This dissertation consists of three essays in empirical asset pricing concerning how investor attention and social interactions impact information diffusion in financial markets. In the first essay, I investigate how different investor attention facilitates information diffusion through the customer-supplier network. I find that retail attention improves information incorporation into asset prices and plays a stabilizing role in financial markets. In contrast, institutional attention exhibits a diminished role if I control for retail attention. In addition, I show that the attention of a potential group of informed retail investors, local retail investors, plays a price-stabilizing role beyond that of uninformed (non-local) retail investors. My results provide a refinement on the view of the role played by retail investors. While the literature argues that retail investors destabilize prices, my findings suggest that at least a group of informed retail investors can stabilize financial markets.

In the second essay, I examine how social connectedness affects fund manager stock holdings during the COVID19 pandemic. I exploit the recent outbreak of COVID-19 as an exogenous shock to people's beliefs on the future economic condition to examine how fund managers from different regions make different decisions on stock holdings. By applying a unique dataset from Facebook that measures social interaction among different regions, I am able to identify managers from COVID-19 hotspot counties and those highly socially connected to the hotspot counties. I am also able to identify fund managers that are skilled using standard methodologies exploiting fund alpha and other performance metrics. The results show that managers located in or socially connected to hotspot counties sold more stock holdings during the outbreak of COVID-19 in the first quarter of 2020. However, such reductions appear panic-driven given subsequent behavior and outcomes and

in particular given the contrasting behavior and outcomes for skilled versus unskilled fund managers. The evidence suggests that social interaction can intensify salience bias even for institutional investors if they are unskilled, but skilled managers appear relatively impervious to the deleterious effect of social networking.

Finally, in the third essay, I explore the role of institutional and retail attention in the context of the media news releases and find nuanced evidence of the costs and benefits to market price adjustments flowing from investor attention to news. I show that retail attention does indeed destabilize financial markets by inducing price overreactions to positive news, but only if it is from uninformed retail investors. I find that when retail attention destabilizes the market, it is when retail investors appear to struggle digesting complex business information and then only if the news is of a positive sentiment; negative sentiment news and retail investor attention are not associated with market instability, possibly a result of the well-known reluctance of retail investors to short sell. I also find that institutional attention plays a stabilizing role in any context I explore, complex or simple news, positive or negative news sentiment, with or without retail investor attention.

ACKNOWLEDGEMENTS

My greatest thanks should be given to my respected supervisor, Dr. Mark Kamstra. This thesis would not have been accomplished without his continuous supports and guidance. Besides the detailed support during my PhD journey, he has been a tremendous mentor for me with the goal of becoming an independent researcher and a righteous person. Words cannot express how grateful I am to him.

I also would like to express my appreciation to my committee member: Dr. Ming Dong, who provided not only advices regarding academic training, but also mental supports during the hardship throughout this journey. I will not forget his timely and precious support. Special thanks should also go to my committee member: Dr. Kee-Hong Bae. His advice regarding my career path that he offered from his own experience and the scarce information regarding academic support are very meaningful to me. And I am also grateful to Dr. Lin Peng and D. Joann Jasiak for being my external and internal-external examiner. Their suggestions and comments greatly enhance the quality of this thesis.

I give my sincere thanks to Shiu-Yik Au, Nazzanin Babolmorad, You-Ta Chuang, Raymond Kan, and Nadia Massoud (in alphabetical order), who provided valuable suggestions for me over the past few years. I also give appreciation to all other faculty members and fellow PhD students at York University for their thoughtful comments on this work.

Finally, I want to give my sincere appreciation to my beloved family, especially to Guoqing Zhou, Lixia Wang, Dongxiang Yin, Guolei Zhang, Qun Zhang, Yike Liu, and John Zhou for letting me concentrate on thesis work, and to my friends Carl Bezanson, Brian Ji, Congsong Zhang, and Janice Zhang for their generous mental supports.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Whose Attention Matters? Evidence from the Return Predictability between Economically Linked Firms

1.1 Introduction

Investor attention, as a limited cognitive resource (Kahneman, 1973), has been recognized as an important source of anomaly returns in financial markets. Traditional asset-pricing theories assume that information will be incorporated into asset prices instantaneously, but this could only happen if it attracts investors' attention. Psychological research shows that people can only process a limited amount of information and, as a result, are subject to limited attention constraints. There is an extensive literature attributing anomaly returns resulting from slow information diffusion to investor inattention (e.g. Hirshleifer and Teo, 2003; Hirshleifer, Hou, Teoh, and Zhang, 2004; Hou, 2007; Cohen and Frazzini, 2008; Hirshleifer, Lim, and Teoh, 2011; among others).

Recently, an emerging literature discusses how different types of investor attention, such as institutional and retail attention, may affect asset prices. The impact of institutional attention is unanimous. Evidence in previous studies suggests that institutional attention facilitates information incorporation into asset prices. (Ben-Rephael, Da and Israelson, 2017; Ben-Rephael, Carlin, Da, and Israelsen, 2020; Liu, Peng, and Tang, 2019; Da, Hua, Hung, and Peng, 2020; among others). In contrast, the impact of retail attention is ambiguous. On the one hand, voluminous studies criticize retail investors for being subject to behavioral biases and show that retail attention deviates price away from the fundamental value (Barber and Odean, 2008; Da, Engelberg, and Gao, 2011; Birru, Chague, De-Losso, and Giovannetti, 2020; among others). On the other hand, some recent works suggest that retail investors, especially those who possess an information advantage, are informed traders (Kaniel, Saar, and Titman., 2012; Kelly and Tetlock, 2017; Boehmer, Jones, Zhang and Zhang, 2020; among others), and the attention paid by informed retail investors, such as local retail investors who may possess private information, is associated with return-anticipating price movements (Cziraki, Mondria, and Wu, 2019).

Understanding the impacts of retail attention on information diffusion in financial markets is important. Although anecdotal evidence suggests that institutional investors manage the majority of financial assets in U.S. financial markets, there is still a significant portion of investment directly managed by retail investors. In addition, retail attention may be very influential for the stocks they prefer to trade but are ignored by institutional investors, provided that there is a large cross-sectional variation of holdings by retail investors. The number of these retail investor-preferred stocks may be quite large because these stocks are likely to be small-cap or micro-cap stocks, which account for a significant portion of the financial market. For example, in December 2017, stocks with a market capitalization less than $2 billion accounts for 6.9% of the total market capitalization, and for 75.6% of the number of stocks in CRSP universe.

In this study, I investigate how return persistence varies with retail attention and I provide further evidence on the comparison of the impact of retail and institutional attention. The analysis is based on a well-known return anomaly – "Customer Momentum," which is first documented in Cohen and Frazzini (2008). They show that the stock return of a firm's major customer predicts the firm's future stock returns. That is, the returns of customer firms, like Walmart, lead the returns of supplier firms in the days, weeks, and months after the customer firm experiences a price shock. A strand of literature attributes this return persistence to the limited attention paid to the customer-supplier relationship by market participants (Cohen and Frazzini, 2008; Guan, Wong, and Zhang, 2015; Cen, Danesh, Ornthanalai, and Zhao, 2015; Agarwal, Leugn, Koana, and Kumar, 2016; Cen, Hertzel, and Schiller, 2019; among others). This anomaly is an ideal testbed for the analysis of retail attention for the following reasons: First, this anomaly is documented to be directly related to investor attention; Second, the customer-supplier relationship is salient public information and is appropriate for the investigation of retail attention because it neither requires a professional background to understand nor requires an advantageous information channel to acquire. Third, the impact of investor attention on information diffusion in this setting of customer-supplier relationships has widespread asset pricing implications. The economic linkage between customer and supplier firms is the cornerstone of the financial market. For instance, customer (supplier) firms account for 51.6% (17.0%) of total market capitalization in CRSP universe in the sample between 2004 and 2017.

Following the recent literature on investor attention, I measure firm-level retail attention using investors' search activity. Specifically, Da, Engelberg, and Gao (2011) establish that the Search Volume Index (SVI) from Google Trends, which measures the search activities of Google users, is a valid measure for the attention of retail investors. There is strong evidence that links this measure to retail investor trading, so that we can identify price stabilizing or de-stabilizing impacts from retail trading by looking at retail (in)attention. See, for instance, Da, Engelberg, and Gao (2011) and Cziraki, Mondria, and Wu, (2019), who establish that retail attention, as measured by Google SVI, is associated with retail investor trading. An advantage of using the search activity-based measure is that it captures investors' demand for information and ensures that investors are indeed paying attention. Measures of the supply of information, such as media coverage, are less direct because investors may not pay attention to the media coverage. To make a comparison between different types of attention, I also need to measure institutional attention. Here, two measures are available from the literature. One is mutual fund common ownership suggested by Cohen and Frazzini (2008), which measures the extent to which mutual funds commonly own both customer and supplier firms. The intuition is that mutual fund managers who hold both customer and supplier firms are more likely to be attentive to both firms, and the information diffusion between these firms is expected to be faster. Another institutional attention measure is suggested by Ben-Rephael, Da, and Israelson (2017), who propose the Daily Max Readership index from Bloomberg terminals. This index captures Bloomberg users' search and reading activities. Ben-Rephael, Da, and Israelson (2017) argue that it is a valid measure for institutional attention because the users are primarily professionals in the financial market. Note that both Google and Bloomberg measure capture investors' search activity, and they provide a comparison between retail and institutional attention.

To investigate whether attention affects the return persistence between customer and supplier firms, I apply two standard asset pricing tools: Long/Short trading strategy and multiple regressions in the spirit of Fama and Macbeth (1973). The "customer momentum" L/S trading strategy is implemented by buying the firms whose major customers yield the highest returns (top quintile) and short selling the firms whose major customers yield the lowest returns (bottom quintile) in the previous month. To explore whether limited attention gives rise to the return persistence, I first split the sample into high and low attention groups, then compare the abnormal return between groups. The abnormal return controls for several well-known risk factors documented in Fama and French (1993), Fama and French (2015), Carhart (1997), and Pastor and Stambaugh (2003). The limited attention hypothesis suggests that the abnormal return will be higher in the low attention group than in the high attention group, if a specific type of attention facilitates the information diffusion between customer and supplier firms. As a robustness test, I apply multiple regressions in the spirit of Fama and Macbeth (1973) to tease out the marginal effect of the main variable, investor attention, while controlling for other well-known factors documented in the literature.

The main finding of this study suggests that retail attention attenuates return persistence from customer to supplier firms. Specifically, in the group with low retail attention, there is both an economically and statistically significant “customer momentum” abnormal return. On the contrary, the abnormal return in the group of high retail attention is nearly zero in magnitude and statistically insignificant. This finding is consistent with the limited attention hypothesis for retail investors. That is, when retail investors are subject to attention constraints, a shock on customer firms does not fully transmit to supplier firms, resulting in a return persistence. This finding also suggests that the return persistence disappears when retail investors pay attention.

Another finding is that the association between return persistence and institutional attention is weak. Using mutual fund common ownership as a measure, I find similar results as those reported by Cohen and Frazzini (2008), in that the abnormal return of L/S trading strategy is statistically significant in the subsample with the implied lower institutional attention (as measured by common ownership of both customer and supplier firms). Similarly, the abnormal return in the low institutional attention group, measured based on the Daily Max Readership from Bloomberg Terminal, is marginally statistically significant. However, I find that the effects of institutional attention are dominated by that of retail attention. Specifically, I apply a double sort approach by first sorting the sample into high and low retail attention groups, then sorting each retail attention group into high and low institutional attention groups. The results show that the abnormal return in low retail attention group is large and mostly statistically significant regardless the level of the institutional attention. Meanwhile, the abnormal return does not have significant a difference between different levels of institutional attention groups within each level of retail attention.

To further delve into these phenomena, I investigate the effects of attention in different samples of firm size, exploring if retail attention measure is merely capturing small and micro-cap firm effects. I find that return persistence is largely isolated to small firms, but the core result of my study on retail attention remains; the (small) firms that suffer from a lack of retail investor attention exhibit return persistence. Interestingly, institutional investors appear to play no part in this. These results can be understood in the framework of Chuprinin, Gorbenko, and Kang (2019). They propose a rational equilibrium model to solve the optimal attention allocation problem when institutional investors aim to maximize their profits through arbitrage on mispriced stocks. Their model shows that institutional investors, especially those with more capital, prefer to devote more attention to investigate the mispricing in larger firms because firm size determines the total arbitrage profits they can extract. In other words, the lower arbitrage profits for smaller firms may be less attractive to institutional investors, suggesting that institutional attention will play a secondary role in the context of return persistence for smaller firms.

In addition, I explore whether the attention of local retail investors, which are documented as informed investors by previous studies (Ivkovic and Weisbernner, 2005; Massa and Simonov, 2006; among others), facilitates the information diffusion from customer to supplier firms. This is achieved by comparing the return persistence between high and low attention group of local retail investors. I define local retail investors as those who reside in the same state as the location of the supplier firms' headquarters. As a placebo test, I also explore the impact of non-local retail investor attention, retail investors who would have no local knowledge of the supplier firms. Non-local retail investors are defined as retail investors who reside in the same state as the location of the customer firms' headquarters.

The results on local retail attention suggest that the attention of local retail investors facilitate information incorporation from customer to supplier firms. Specifically, I find that the abnormal return of “customer momentum” strategy is economically and statistically significant in the group of low local retail attention group but is not statistically significant in the high local retail attention group. In contrast, the abnormal returns in high and low nonlocal retail attention groups are similar.

The finding on local retail attention also implies a mechanism through which retail investors’ attention may facilitate information incorporation. In the context of this study, a high level of local retail attention on firms may reflect local investors’ observation of shocks on customer firms either from public or private channels. Then their active involvement would facilitate the transmission of this shock to supplier firms. This mechanism is analogous to the one proposed by the conceptual framework in Cziraki, Mondria and Wu (2019), which shows that once local retail investors observe private signals, it is optimal for them to collect more information, such as searching public information through Google, to confirm the signal. Furthermore, Cziraki, Mondria and Wu (2019) provides empirical evidence suggesting that local retail investors are informed, and their attention predicts future returns of local firms.

Finally, I perform a series of robustness tests to ensure that my results are not capturing effects already documented in the literature. Specifically, I first verify the results using multiple regression in the spirit of Fama and Macbeth (1973), controlling for other well-documented factors that could be alternative explanations for the return predictability between customer and supplier firms, such as industry momentum effects (Moskowitz and Grinblatt, 1999), cross-industry momentum effects (Menzly and Ozbas, 2010), customer firms’ momentum effects, and supplier firms’ own momentum effects (Jegadeesh and Titman, 1993). Moreover, I verify the robustness of the retail attention effects after controlling for other attention measures, such as institutional attention (Ben-Rephael, Da, and Israelson, 2017), mutual fund common ownership (Cohen and Frazzini, 2008) and firm size. In addition, I also construct long/short trading strategy using the idiosyncratic return suggested by Burt and Hrdlicka (2016) to rule out the possibility that the "customer momentum" strategy applied in this study are biased due to a misspecified asset pricing model. Furthermore, I apply double sort approach to explore whether retail attention simply picks up firm size effects and find that both firm size and retail investors attention matters in the case of return predictability. In other words, the main effect of return predictability is concentrated on smaller firms with low attention from retail investors.

The rest of the paper is organized as follows. Section 1.2 provides a discussion about the related literature. Section 1.3 provides a detailed description of the data sources and outlines the empirical methodology for the measurement of different types of investor attention. Then in section 1.4, I provide the main findings regarding the "customer momentum" trading strategy and the multiple regressions in the spirit of Fama-Macbeth (1973). Section 1.5 describe different kinds of robustness test. Finally, I discuss implications of the impacts of retail attention in section 1.6 and conclude in section 1.7.

1.2 Related Literature

There is a broad literature discussing the impact of limited investor attention on delayed information diffusion theoretically and empirically. Theoretical models relate investor inattention to price underreaction due to different reasons, such as market segmentation (Merton, 1987), alternative presentation of accounting information (Hirshleifer and Toeh, 2003), limited attention on learning behavior (Peng and Xiong, 2006), and neglecting accrual accounting information

(Hirshleifer, Lim and Teoh, 2011). A growing body of literature documents empirical evidence of underreaction to information driven by limited attention from different sources, such as inattention to customer-supplier relationships (Cohen and Frazzini, 2008; Cen, Hertzel, and Schiller, 2019), inattention to earning news due to distraction to a large number of earning announcements (Hirshleifer, Lim, and Teoh, 2009), inattention to earnings announcement on Friday, close to weekend (DellaVigna and Pollet, 2009), and inattention of retail investors to earning news due to distraction of macroeconomic news (Liu, Peng, and Tang, 2019).

An emerging empirical literature is concerned with the different impacts of attention on financial markets from retail and institutional investors. One strand of literature highlights the price-stabilizing role of institutional investors or the price-destabilizing role of retail investors. For example, Barber and Odean (2008) find that attention-grabbing events induce buying pressure from retail investors, not institutional investors. Ben-Rephael, Da, and Israelson (2017) show that the post-announcement drifts caused by earning announcement and analyst recommendation change weaken conditional on abnormal attention of institutional, not retail investors. Ben-Rephael, Carlin, Da, and Israelsen (2018) provide evidence that only abnormal institutional attention facilitates price discovery before the filing period of SEC 8-K filings. Chuprinin, Gorbenko, and Kang (2019) document the evidence that abnormal institutional attention improves price correction of mispricing at earning announcement. Da, Hua, Hung, and Peng (2020) find that aggregate firm-level retail attention negatively predicts the market return whereas aggregate institutional attention weakly but positively predicts the market return. They argue that retail attention slows down the incorporation of negative news to the market due to behavioural biases such as disposition effect (Shefrin and Statman, 1985) and ostrich effect (Galai and Sade, 2006).

Another strand of literature concerns the stabilizing role retail attention plays on market efficiency and is most closely related to this study. Liu, Peng, and Tang (2019) find that retail inattention results in lower contemporaneous return response to earning announcement even conditioning on abnormal institutional attention. They use the announcement of important macroeconomic news as an exogenous shock on retail attention to investigate the return response to earning news. Song (2020) looks at the role of retail attention to accounting information during earning announcement period. Consistent with the results in Liu, Peng, and Tang (2019), she finds stronger contemporaneous return reactions and weaker post-announcement drift on earning news when retail investors pay abnormal attention to accounting information. In my study, I investigate the role of retail inattention on return persistence between customer-supplier relationships. In this context, the focal firms (supplier firms) are mostly smaller firms that are not predominantly owned by institutional investors. So, the investigation on these firms is appropriate for teasing out the impact of retail attention largely free of the impact of institutional attention. Moreover, I leverage the well-documented "home bias" phenomenon to investigate the impact of local retail attention and the location-based information bias.

This study also relates to the literature on the limited attention hypothesis concerning the return predictability between customer and supplier firms. The seminal work of Cohen and Frazzini (2008) first document this return predictability and find that institutional inattention is associated to the return predictability. Cen, Hertzel, and Schiller (2019) propose a novel measure of information diffusion speed between customers and suppliers and show that the speed is associated with investor attention, proxied by active institutional crossholding and analyst dual coverage. Agarwal, Leung, Konana, and Kumar (2017) show that it is less likely to observe return predictability when investors pay co-attention, measured using a co-viewing pattern of firms on

Yahoo! Finance. My study extends this literature by documenting the association between the return predictability and retail inattention, controlling for institutional attention.

There is also an interesting debate concerning the role played by retail investors. Historically, retail investors are criticized as noise traders or traders subject to behavioral bias, such as Barber and Odean (2000), Kumar and Lee (2006), Hvidkjaer (2008), and Barber, Odean, and Zhu (2009). However, recent studies document evidence showing that some retail investors can also be informed traders, such as Kaniel, Saar, and Titman (2008), Kelley and Tetlock (2013, 2017) and Boehmer, Jones, Zhang, and Zhang (2020). My paper expands on the evidence of price-stabilizing behavior of retail investors, at least in some important settings.

1.3 Data

The data involved in this study come from several sources. The retail attention data come from Google Trends; institutional attention data come from Bloomberg; mutual fund common ownership data come from Thomson Reuters Mutual Fund Holdings (S12); the customer-supplier relationship data come from Compustat Customer Segment File; the stock return information come from the Center for Research in Securities Prices (CRSP) and accounting information comes from Compustat. My analysis is primarily conducted at monthly frequency between June 2004 and May 2017. I choose 2004 as the starting point because it is the year Google Trends becomes available. My sample consists of all firms in CRSP/Compustat universe that are: i) identified with at least a customer-supplier relationship; ii) common equities with share code of 10 or 11 in CRSP; iii) not missing market equity or book equity values; iv) with a closing price above or equal \$5 at the end of a month.

1.3.1 Retail Investor's Attention Data

I obtain the retail investor attention from Google Trends, the Search Volume Index (SVI), which aggregates geographically specific search activities by keyword. This is a time-series data set, which is normalized between zero and one hundred and is constructed on a daily, weekly, and monthly basis. To include as much information as possible, I exploit the daily SVI for each firm involved in a customer-supplier relationship.

Following Da, Engelberg, and Gao (2011) and numerous studies in the literature, I choose the ticker symbol as the keyword to represent the attention of retail investors with respect to a firm. However, to further improve the data reliability, I apply a three-step procedure to tease out the search activities that are made by investors to look for stock related information. First, I choose the "investing" subcategory in a function now provided by Google Trends, which classifies search activity into various categories using proprietary algorithms. As a verification, I double-check top related topics and top related queries that come with the same keyword simultaneously and keep the Google SVI data if top related topics contain "stock" topic or top related queries contain stock related information. Second, for the keywords that fail to pass the verification test or return missing values due to insufficient observations, I choose "all categories" instead of the "investing" subcategory using the same keyword. I also conduct a verification test by looking at the top related topics and top related queries exploring whether the SVI captures search activities related to investing purposes. Third, in the case that a ticker symbol fails to pass the verification tests in previous steps, I augment the ticker symbol with "stock" as the new keyword. I discard firms that cannot be verified as reliable results.

It is important to apply the three-step procedure to make the Google SVI data largely free of the impact of search activities unrelated to investing. Firm managers have an incentive to design a memorable ticker symbol (Andersen and Larkin, 2018), which could be a common English term (CAT for Catepillar Inc.), English letter (C for CIT group), or well-known abbreviations (GPS for Gap Inc.). Using ticker symbols without treatment would sweep in unrelated searches for many firms. As an illustration, I compare the SVI from Google Trends on keyword - “CAT” based on “All categories” and the “Investing” subcategory and provide the output in Panel A and B of Figure 1.1.

Panel A of Figure 1.1 illustrates the output of a Google Trends search for “CAT” from “All categories”. To be specific, I restrict the SVI to include only the searches made in the U.S., as a ticker symbol may represent different companies in other countries. The time scope is chosen to start from 2004 to 2017, to correspond with the full sample period of my study. In the following, I use default options for the rest of the settings such as “All categories” for category and “Web Search” for the search type. The blue trajectories, labeled as “interest over time,” are the SVI representing the search intensity over time for the keyword “CAT”. Next, Google Trends provides a geographic comparison in the section labeled as “interest by subregion” which indicates that searches for “CAT” come from every region of the United States. At the bottom, an important function of the analysis of the searches provides the related topics and related queries. To sum up, Panel A of Figure 1.1 demonstrates a generic output about the ticker symbol “CAT”. The analysis of the search shows that the top related topics are “Cat – Animal”, and the top related queries are “cats”, both suggesting that this SVI contains mostly the search activities related to the interest in cat, the animal, rather than the Caterpillar Inc., the public company.

As a comparison, Panel B of Figure 1.1 illustrates the graphical output for “CAT” by specifying the “Investing” subcategory. Holding the rest of the settings unchanged, I find that the refined SVI appears to be primarily related to investment queries. Firstly, the region most

Figure 1. 1: Panel A - Google Trends search for "cat" in "All categories"

This Figure provides the illustration of Google Trends search result. Panel A shows the output for a search of "cat" in Google Trend. The region is set to be "United States"; the time range is between June, 2004 and May, 2017; the category is by default chosen as "All categories"; and the search content is "Web Search". The blue trajectories labeled as "Interest over time" plots the SVI for "cat". The map labeled as "Interest by subregion" gives the allocation of search activities over different states. In the bottom, the related topics and related queries show the top rank of the topics and queries throughout the chosen time range.

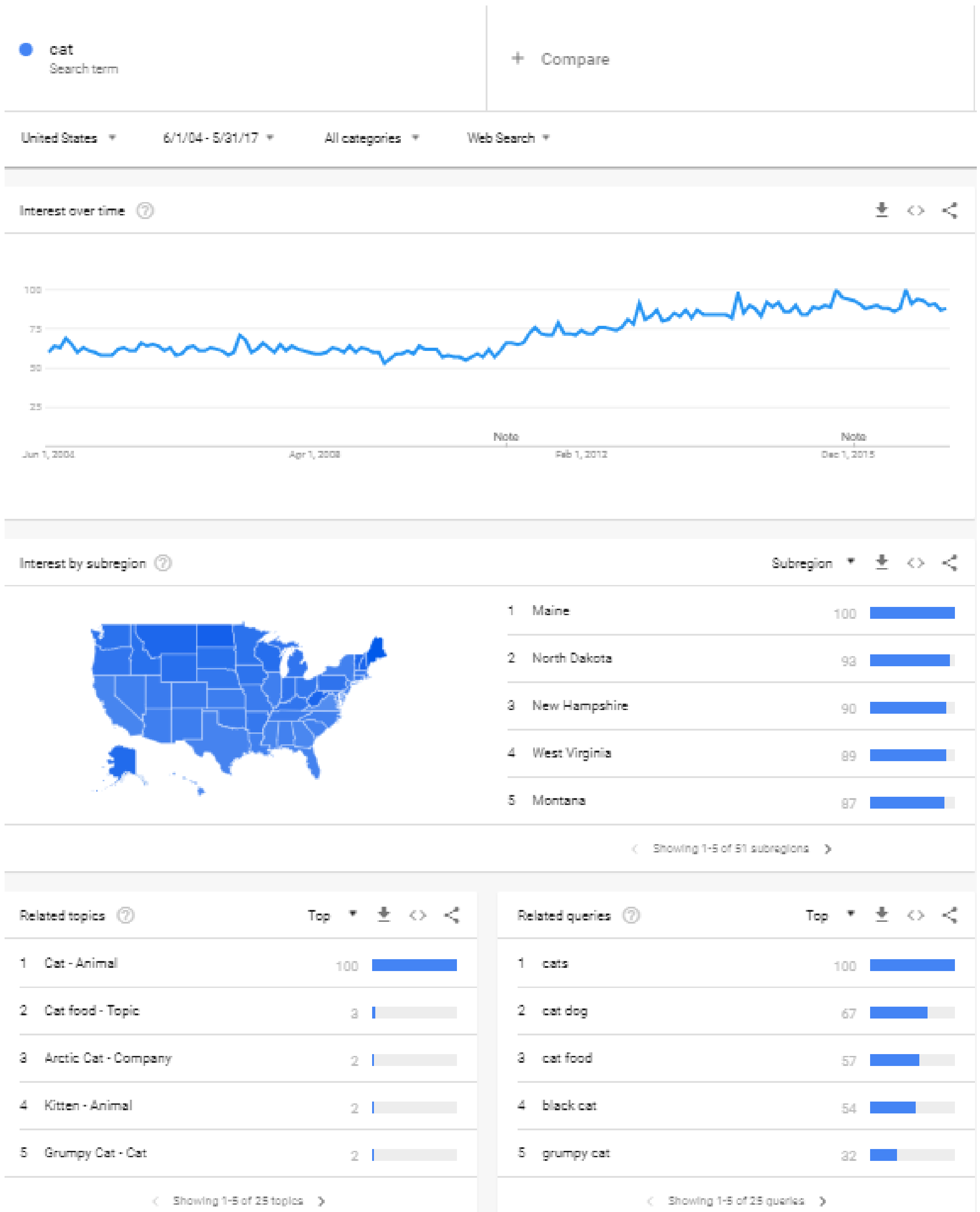
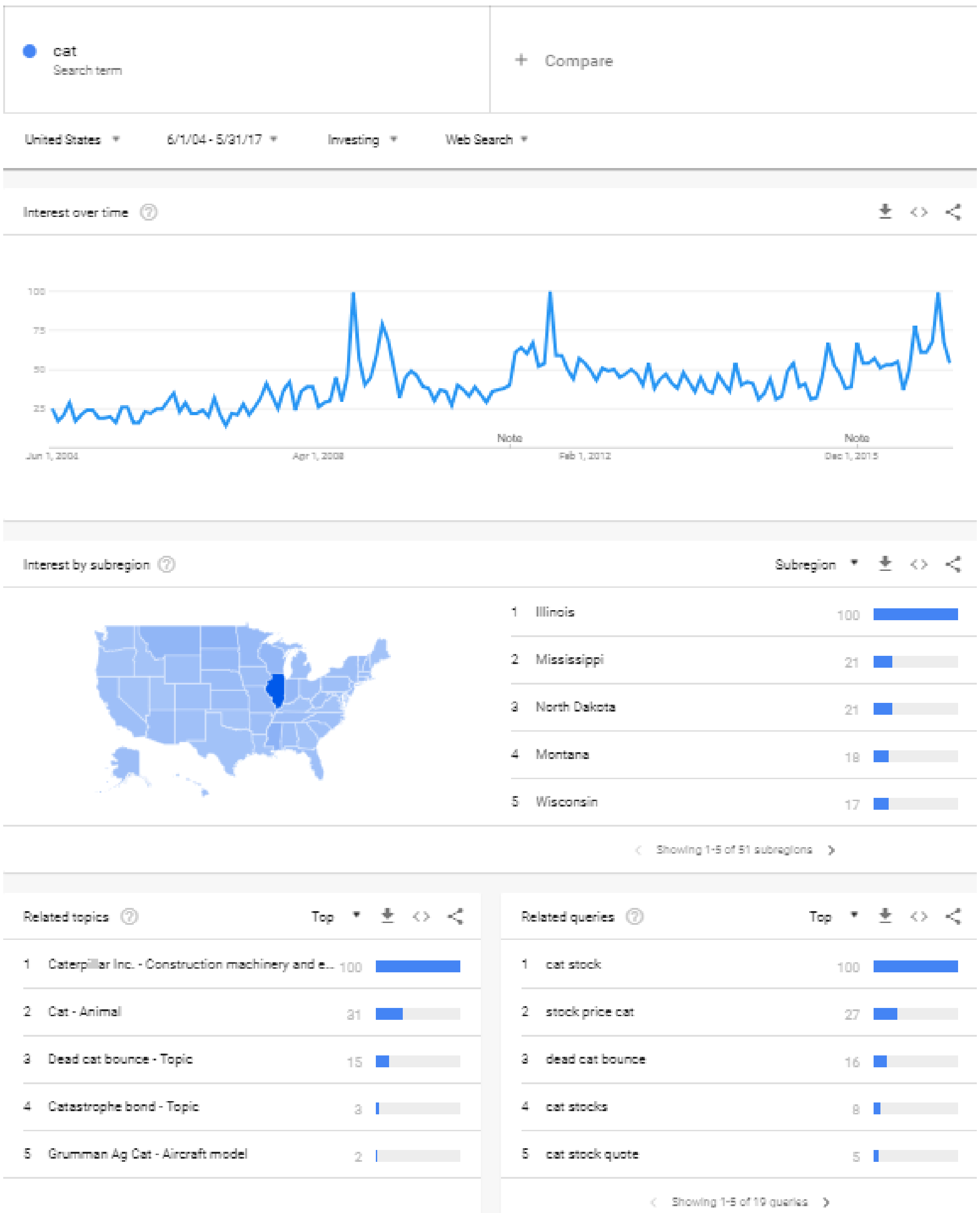


Figure 1. 1: Panel B - Google Trends search for "cat" in "Investing" subcategory

This Figure provides the illustration of Google Trends search result. Panel B shows the output for a search of "cat" in Google Trend. The region is set to be "United States"; the time range is between June, 2004 and May, 2017; the category is specified as "Investing"; and the search content is "Web Search". The blue trajectories labeled as "Interest over time" plots the SVI for "cat". The map labeled as "Interest by subregion" gives the allocation of search activities over different states. In the bottom, the related topics and related queries show the top rank of the topics and queries throughout the chosen time range.



interested in this ticker is “Illinois,” which is the state in which Caterpillar Inc. is headquartered. Additionally, the exact company name represented by the ticker symbol is on the top of the related topics list[1]. The blue bar on the right-hand side shows that the number of searches identified as related to the company is four times as high as those related to the “Cat- Animal”, the primary topic unrelated to the firm. Last, the top related queries are a combination of the ticker symbol and additional information such as “stock”, “stocks”, “stock quotes”, and “stock price” which confirms that the results from the “Investing” subcategory are indeed related to search activities made by investors.

Finally, I manually check the ticker symbols that are common English terms or well-known abbreviations and verify whether my rule of validation works. It turns out that the three-step procedure successfully identifies the investing related searches in most of the cases except for 16 firms out of my total of 2475 firms. I delete these firms from the total sample, and the list of these firms is available upon request.

1.3.2 Local Retail Investor’s Attention Data

I obtain the local retail attention data using a special feature of Google Trends that provides the intensity of search activities from specific regions. By specifying a geographic location, such as a state, one can obtain the Search intensity from that state. Using previous example about ticker symbol “CAT”, I provide the output in panel C and D of Figure 1.1 about the SVI from two different states that represent customer and supplier firm’s local retail investors’ attention respectively. In specific, Panel C of Figure 1.1 illustrates the graphical output for “CAT” by

[1] Basistha, Kurov, and Wolfe (2018) suggested to use the topic categories, related queries, and regions to find the search activity data that is more related to the investor attention.

Figure 1. 1: Panel C - Google Trends search for "cat" in Illinois

This Figure provides the illustration of Google Trends search result. Panel C shows the output for a search of "cat" in Google Trend. The region is set to be "Illinois"; the time range is between June, 2004 and May, 2017; the category is specified as "Investing"; and the search content is "Web Search". The blue trajectories labeled as "Interest over time" plots the SVI for "cat". The map labeled as "Interest by subregion" gives the allocation of search activities over different states. In the bottom, the related topics and related queries show the top rank of the topics and queries throughout the chosen time range.

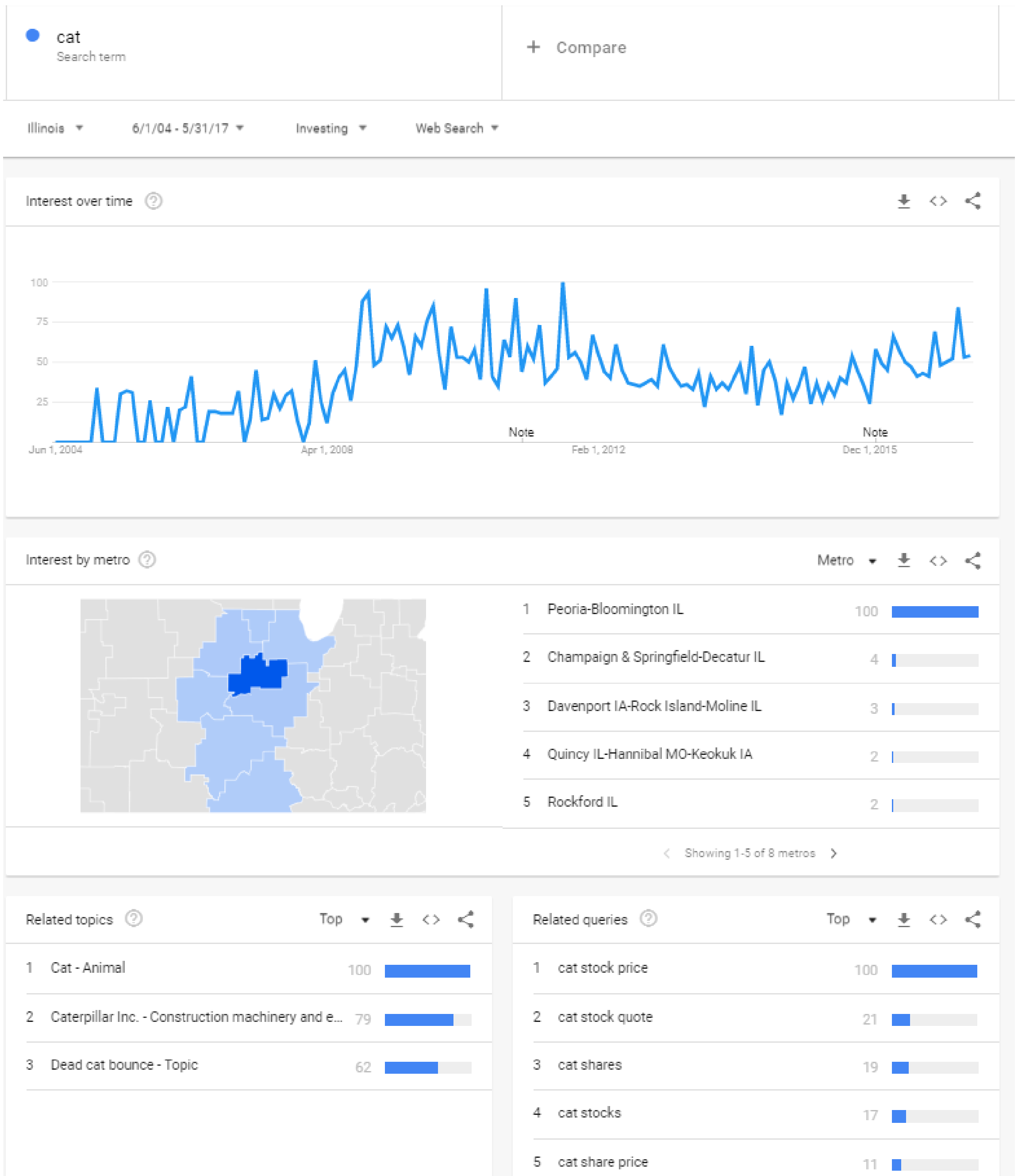
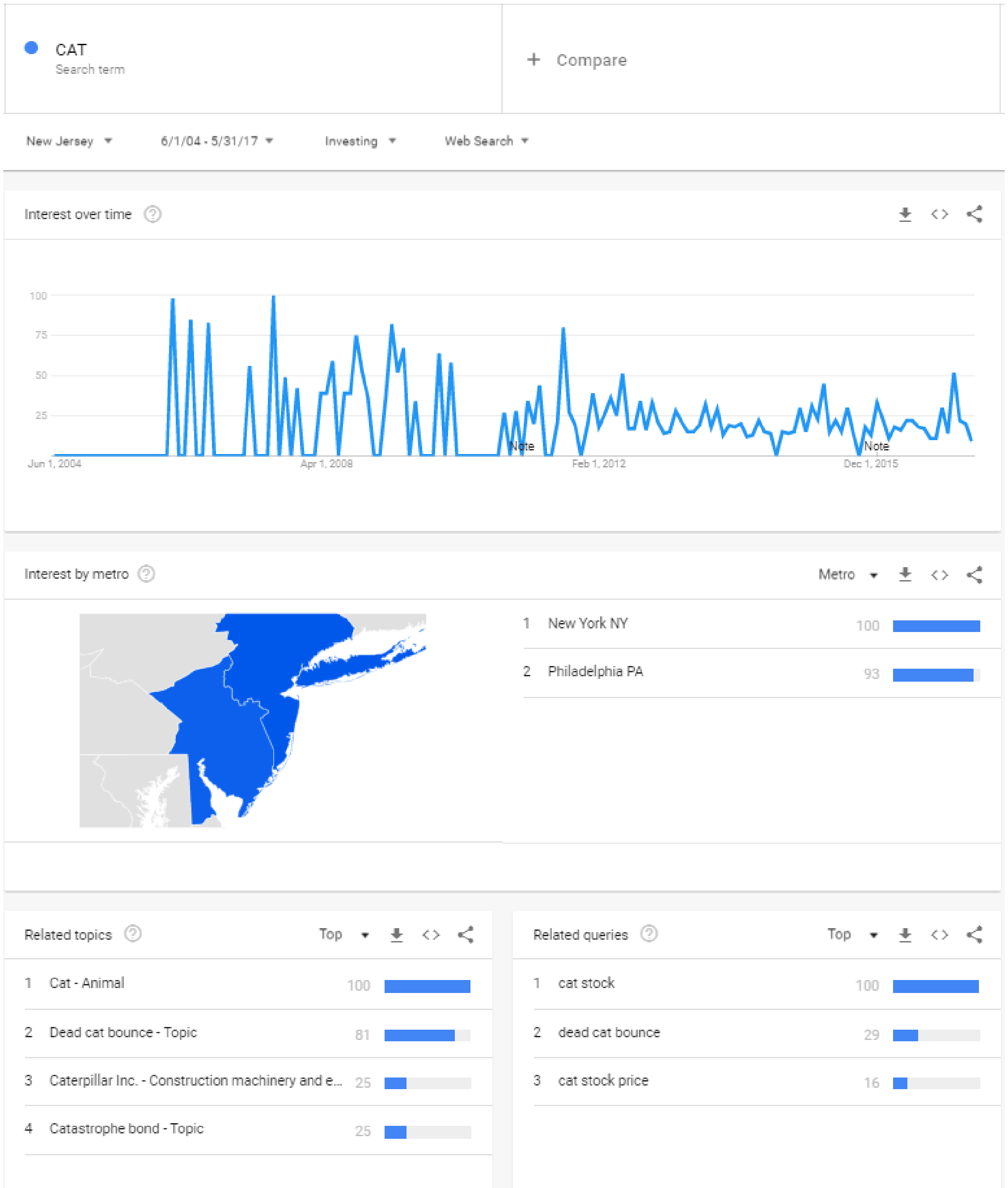


Figure 1.1: Panel D - Google Trends search for "cat" in New Jersey

This Figure provides the illustration of Google Trends search result. Panel D shows the output for a search of "cat" in Google Trend. The region is set to be "New Jersey"; the time range is between June, 2004 and May, 2017; the category is specified as "Investing"; and the search content is "Web Search". The blue trajectories labeled as "Interest over time" plots the SVI for "cat". The map labeled as "Interest by subregion" gives the allocation of search activities over different states. In the bottom, the related topics and related queries show the top rank of the topics and queries throughout the chosen time range.



specifying the "Investing" subcategory in Illinois, in which Caterpillar Inc., the customer firm, is headquartered. Panel D of Figure 1.1 provides the output in New Jersey, in which one of Caterpillar Inc.'s supplier firms, Orbcomm Inc., is headquartered. I also apply "investing" category as the filter to reduce the noise in the data that may related to non-investing purpose. The top related topic and top related queries confirms that the obtained SVI are related to investing purpose. Note that while the SVI in Panel B, C, and D capture some common spikes, such as the spike in Aug, 2011, and Feb, 2017, they also exhibit unique patterns, suggesting that each of them captures the attention of different groups of retail investors.

1.3.3 Institutional Investor's Attention Data

Ben-Rephael, Da, and Israelson (2017) propose a measure for institutional investors' attention, which is called News Heat - Daily Max Readership and is constructed based on search and reading activities from Bloomberg Terminal users. Bloomberg tracks the search and reading volume related to a specific stock by its users. To give more weight to the behavior of users actively seeking the news of a given firm, Bloomberg assigns a value of 10 to the search attempt, and 1 for simply reading news articles about a specific firm. Then it aggregates all these activities about a specific firm each hour and compares the average hourly count for the previous 8 hours to the average of the total hourly count for the previous 30 days. After the comparison, Bloomberg assigns a score between 0 and 4 to identify how active the most recent 8 hours were compared to the previous days. A value of 0 signals that the recent average hourly search and reading intensity is below 80th percentile of hourly counts for the previous 30 days. A value of 1,2,3, or 4 signals that the recent average is between 80th -90th, 90th -94th, 94th -96th, and greater than 96th percentile of the previous 30 days activity respectively. In the end, for each day, Bloomberg chooses the maximum value for the hourly score to represent the daily search and reading activity.

Ben-Rephael, Da, and Israelson (2017) define an indicator variable that equals 1 when it is above the 94th percentile and 0 otherwise. Readers are directed to Rephael, Da, and Israelson (2017) for a more detailed explanation of the data construction.

1.3.4 Measure of Abnormal Attention Data

Following Ben-Rephael, Da, and Israelson (2017), I define an indicator variable that equals 1 when the search and reading intensity is greater than the 94th percentile as daily institutional investors abnormal attention. I define daily retail investor's abnormal attention based on Google SVI in a similar fashion. I assign a score of 0 or 1 based on the stocks' daily SVI for the previous 30 days compared to the current day's SVI. For example, a score of 0 signals that the current daily SVI is below the 94th percentile of the daily SVI for the previous 30 calendar days.

Since my study focuses on the monthly frequency, I construct the Monthly Retail investors' Abnormal Attention (MRAA) and Monthly Institutional investors' Abnormal Attention (MIAA) by taking the average of daily retail and institutional investors' abnormal attention over the month. As the daily abnormal attention is a dummy variable, the value of monthly average of these dummy variables falls between 0 and 1, reflecting how frequently a firm receives abnormal attention during a certain calendar month. So, a value of 0 means no abnormal attention is paid during the month, and a value of 1 means investors pay abnormal attention on each trading day of the month.

The abnormal attention for local retail investors is constructed in a similar fashion but making use of monthly rather than a daily data .This difference is due to the fact that daily values of Google SVI for a specific region and firm are often missing. To define abnormal attention in monthly frequency, I follow Cziraki, Mondria, and Wu (2019) and take the natural logarithm of the difference between the current monthly SVI and the median level of SVI over the previous three months.

1.3.5 Customer-Supplier Relationship Data

Since 1997, regulation SFAS No. 131 requires public business enterprises in the U.S. to publicly disclose the identify of their customer firms that account for more than 10% of the firm's annual sales. The customer-supplier linkage data in this study is constructed in two steps. Firstly, the name of the customer firms is obtained from the Compustat Customer Segments File, then it is manually matched to the name of the firms in the CRSP/Compustat universe. To make sure the matching is reliable with best efforts, two algorithms are developed independently, and any inconsistent matches between the two algorithms are defined as suspicious matches and are manually scrutinized[2]. Similar hand match approaches are widely used in the literature, such as in Banerjee, Dasguapta, and Kim (2008), Cohen and Frazzni (2008), and Cen, Hertzel, and Schiller (2019).

Following the convention in the literature, I define the valid period for a customer-supplier relationship as starting from six months after the reporting fiscal year-end dates, which allows the market to have enough time to be aware of the relationship and the related accounting information.

1.3.6 Summary Statistics

Table 1.1 provides summary statistics for the data used in this study. The sample period spans 13 years between June 1st, 2004, and May 31st, 2017. I choose 2004 as the starting point because it is when retail attention data first becomes available. Overall, there are 1,628 suppliers

[2] Thanks, Donghyun Kim and Rui Dai for sharing the customer-supplier link data.

and 773 customers, 3,831 unique customer-supplier links, and 146,630 customer-supplier-pair-month observations throughout the sample.

Panel A of Table 1.1 show the firm characteristics for customer and supplier firms, respectively. On average, the market capitalization of customer firms is more than twelve times as large as that of supplier firms. The monthly average combined market capitalization for customer firms is 51.6% of the total market capitalization of the entire market in the CRSP stock universe, though the total number of customer firms accounts for only 6.8% of the total number of CRSP firms. The combined market capitalization of supplier firms is around 17% of the entire market, while the total number of supplier firms is around 13% of the total number of firms[3]. On average, a supplier firm reported 2.4 customer firms on average, and a customer firm corresponds to 5.0 supplier firms. The percentage of sales to customers, on average, is 18.4% of the total sales of the supplier firms, which confirms that the customer-supplier firm pairs have a strong connection in real business operations.

Panel B of Table 1.1 provides the summary statistics for attention measures and firm characteristics used in this study. On average, the Monthly Retail investors' Abnormal Attention (MRAA) on customer firms is 5.8%, suggesting that the incidence of experiencing daily retail abnormal attention is 1.22 times per month (5.8% x 21 working days=1.22). However, the average Monthly Institutional investors' Abnormal Attention (MIAA) on customer firms is 22.2%, suggesting an average customer firm will experience daily institutional abnormal attention 4.66 times per month (22.2% x 21 working days = 4.662). Note that number of observations for Monthly Institutional investors' Abnormal Attention (MIAA) is around half of the observations for Monthly

[3] These summary statistics are consistent with the numbers reported in Cohen and Frazzini (2008).

Table 1. 1: Summary Statistics

This table provides summary statistics for the entire sample. Panel A shows statistics in month level. *Number (Value) % of CRSP stock universe* provides the monthly average percentage of the combined number (market capitalization) of firms out of CRSP stock universe. *Sale percentage* gives the average percentage of sales to a reported major customer accounted for total sales of a supplier firm. Panel B provides statistics in customer supplier pair-month level. MRAA and MIAA denotes Monthly Retail and Institutional investors' Abnormal Attention as described in section 2.3. MF Common Ownership is the number of mutual funds which cross hold both customer and supplier firms divided by the total number of mutual funds which hold supplier firms. Local(nonlocal) Abnormal attention are the abnormal attention paid by local (nonlocal) retail investors, defined using firm's headquarter location. The units of market capitalization are million of US dollars.

Retail

Variables	Nobs	Mean	Median	Std	P25	P75
Panel A: Time Series (156 Monthly Observations, 2004 - 2017)						
Number % of CRSP stock universe – customer firms		6.8%	7.1%	1.1%	5.9%	7.6%
Value % of CRSP stock universe – customer firms		51.6%	51.8%	3.5%	48.7%	54.1%
Number % of CRSP stock universe – supplier firms		12.8%	13.0%	1.1%	12.3%	13.5%
Value % of CRSP stock universe – supplier firms		16.9%	16.8%	1.4%	16.0%	17.8%
Number % of CRSP stock universe – customer & supplier firms		18.1%	18.6%	1.8%	16.9%	19.6%
Value % of CRSP stock universe – customer & supplier firms		58.4%	58.7%	3.8%	55.5%	61.1%
Number of customers per supplier		2.4	2.0	2.3	1.0	3.0
Number of suppliers per customer		5.0	1.0	11.7	1.0	4.0
Sale percentage		18.4%	14.9%	14.0%	10.6%	21.8%
Panel B: Attention Measures & Firm Characteristics						
MRAA on customer firms	146,630	5.8%	6.5%	4.7%	3.2%	9.7%
MIAA on customer firms	79,385	22.2%	19.1%	17.8%	9.1%	33.3%
MF Common Ownership	146,630	27%	24%	20%	15%	36%
Local Abnormal attention	146,630	0.06	0.00	1.02	-0.13	0.19
Nonlocal Abnormal attention	146,630	0.05	0.00	1.03	-0.20	0.26
Market Cap (Million) - customer firms	146,630	70,031	29,308	89,083	10,141	101,945
Market Cap (Million) - supplier firms	146,630	5,482	878	19,467	339	2,641
Book to Market - customer firms	146,630	0.55	0.38	0.63	0.28	0.62
Book to Market - supplier firms	146,630	0.55	0.44	0.46	0.27	0.71
Return - customer firms	146,630	0.80%	0.87%	8.94%	-3.53%	5.21%
Return - supplier firms	146,630	0.76%	0.56%	12.84%	-5.81%	7.05%

investors' Abnormal Attention (MRAA) because the data in Bloomberg is only available since February 2010. In addition, it is noted that the average value of mutual funds common ownership is around 27%[4]. Lastly, the average supplier and customer firms' local investors abnormal attention is 0.06 and 0.051 respectively. Note that this attention measure is the natural logarithm of the difference between the current monthly SVI and the median level of SVI over the previous three months, with the level of SVI scaled between 0 and 100. Finally, untabulated results show, unsurprisingly, that all the measures of investor attention are positively related to each other, with the strongest correlations between retail and institutional attention to customer firms.

1.4 Empirical Results

Table 1.2 provides the main results of this study. Panel A shows the equal weighted abnormal return of portfolios formed by sorting supplier stocks based on their linked customers' stock performance in the previous month. The supplier stocks are sorted into five quintiles. The abnormal return for each quintile portfolio is presented from Q1 (low) to Q5 (high), which includes the group of suppliers linked to the customers whose return ranked from the bottom to the top quintile in previous month. The rightmost column provides the abnormal return of a zero-cost portfolio that takes a long position in the suppliers with the highest quintile of customer return and a short position in the suppliers with the lowest quintile of customer return. For brevity, I only provide the abnormal return with respect to the Fama and French (2015) five-factor model, augmenting the model with momentum factor (Carhart,1997) and liquidity factor (Pastor and Stambaugh, 2003).

[4] Following Cohen and Frazzini (2008), I define mutual fund common ownership using the number of funds holding both customer and supplier firms divided by the number of funds holding supplier firms only.

Table 1. 2: Abnormal Return (EW) of Customer Momentum Strategy, Sort by Retail Investors' Attention

This table provides the equal weighted abnormal return of customer momentum strategy for calendar-time portfolio for full sample and subsamples split by the level of Monthly Retail Abnormal Attention (MRAA) on customer firms. Panel A provides the abnormal return of the portfolios based on the full sample. Panel B shows the abnormal return in the groups of high and low Monthly Retail Abnormal Attention (MRAA) on customer firms. In each calendar month, supplier firms are sorted based on the return of major customers in the previous month and are grouped into five quintiles. Q1 and Q5 denote the groups corresponding to the lowest and highest customer return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers. The abnormal return (Alpha) is the intercept on a regression of monthly excess return from the rolling strategy. T stat denotes the t-statistics rounded to two decimals. The explanatory variables are monthly returns of Fama and French (1993) and Fama and French (2015) mimicking portfolios, the Carhart (1997) momentum factor, and the Pastor and Stambaugh (2003) liquidity factor. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

		Q1 (Low)	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Panel A: Full Sample, 2004-2017							
	Alpha	-0.280*	-0.178	-0.127	-0.027	0.145	0.425**
	T stat	[-1.94]	[-1.36]	[-1.00]	[-0.21]	[1.05]	[2.20]
	N	105	100	101	100	97	
	Size	5,440	5,673	5,652	5,961	4,570	
Panel B: Split by Monthly Retail Abnormal Attention (MRAA) on Customer Firms, 2004-2017							
Low	Alpha	-0.410**	0.067	-0.101	0.064	0.422**	0.832***
	T stat	[-2.14]	[0.42]	[-0.61]	[0.38]	[2.30]	[3.11]
	N	63	59	58	58	56	
	Size	5,298	5,547	5,266	6,137	4,711	
High	Alpha	-0.305*	-0.373*	-0.121	-0.033	-0.277	0.028
	T stat	[-1.73]	[-1.68]	[-0.60]	[-0.17]	[-1.32]	[0.10]
	N	47	42	42	42	38	
	Size	4,934	6,078	5,051	5,416	3,521	

The result in Panel A shows that the "customer momentum" strategy implemented in my sample yields a statistically significant abnormal return of 0.425% per month (t-statistic = 2.20), which is consistent with the findings of the return predictability with customer-supplier relationship documented in Cohen and Frazzini (2008). The abnormal return on the long leg is 0.145% per month (t-statistic = 1.05), which is statistically insignificant. The abnormal return on the short leg is -0.280% per month (t-statistic = -1.94), which is marginally statistical significant. The lead-lag effects are stronger in the short leg, which implies that the delayed reaction of a stock is stronger for negative news. The abnormal return in value-weighted portfolios is not displayed here and is not statistically different from zero. This suggests that the return predictability is mainly driven by smaller firms in the sample. The average number of supplier stocks in each quintile ranges from 97 to 105, suggesting a sufficient amount of stocks in each quintile. To ensure that the abnormal returns are not driven by illiquid securities, I exclude the stocks with a closing price below $5 at the end of the previous month and include the liquidity factor (Pastor and Stambaugh, 2003). Note that the abnormal returns across quintile portfolios show a monotonic pattern, ranging from -0.280% to 0.145%. In the following step, I take attention measures into consideration and check whether it can improve the profitability of the trading strategy and shed light on the importance of retail versus institutional investor attention.

1.4.1 Split Sample Using Retail Investors' Attention

In the following analysis, I investigate whether the attention of retail investors plays an important role in the return predictability documented above. Panel B of Table 1.2 provides the comparison of the abnormal return from the "customer momentum" strategy between the subsample with low and high retail attention. The suppliers are grouped into two subsamples based on the level of Monthly Retail Abnormal Attention (MRAA) on customer firms in the previous month, which is just the average of daily retail investors' abnormal attention over the previous month. Then, the customer momentum strategy is implemented within each subsample. The result shows that the abnormal return of L/S trading strategy in the low attention group is 0.832% (t-statistic = 3.11), which is almost two times as large as that in the full sample, 0.425%. In other words, when retail investors pay less attention to a customer firm, the return of the customer firm has stronger predictability for the return of the associated supplier firm in the following month. In contrast, when retail investors pay relatively high attention to a customer firm, the abnormal return is 0.028% (t-statistic=0.10), which is almost zero and statiscally insignificant, suggesing little return predictability for the supplier firm. The abnormal return increases almost monotonically in the low retail attention group, while it fluctuates without a clear pattern in the high retail attention group, consistent with return predictability in the low attention group. The average number of firms in low retail attention group is larger than that in high group because some firms clustered at zero level of retail attention, however, the average size of firms, represented by the market capitalization of the firms, is similar between the low and high group, implying that this way of grouping is not capturing the lead-lag effect induced by the firm size. Furthermore, the evidence that the statistically significant abnormal returns in both short and long leg in low attention group suggests that retail investors of suppliers underreact to both negative and positive news on customers.

1.4.2 Split Sample Using Institutional Investors' Attention

Now, to make a comparison for the effects of attention between different types of investors, I repeat the analysis using the attention measure on institutional investors. Panel A of Table 1.3 replicates Panel A of Table 1.2 and is presented for ease of comparison to results in Panels B and C. Panel B of Table 1.3 shows the result of the subsamples with low and high mutual fund common ownership. Cohen and Frazzini (2008) show that the return predictability is stronger in low

Table 1. 3: Abnormal Return (EW) of Customer Momentum Strategy, Sort by Institutional Attention

This table provides the equal weighted abnormal return of customer momentum strategy for calendar-time portfolio for full sample and subsamples split by the level of institutional investors' attention. Panel A provides the abnormal return of the portfolios based on the full sample. Panel B presents abnormal return in the groups of high and low mutual fund common ownership. Panel C shows the abnormal return in the groups of high and low Monthly Institutional Abnormal Attention (MIAA) on customer firms. In each calendar month, supplier firms are sorted based on the return of major customers in the previous month and are grouped into five quintiles. Q1 and Q5 denote the groups corresponding to the lowest and highest customer return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers.The abnormal return (Alpha) is the intercept on a regression of monthly excess return from the rolling strategy. T stat denotes the t-statistics rounded to two decimals. The explanatory variables are monthly returns of Fama and French (1993) and Fama and French (2015) mimicking portfolios, the Carhart (1997) momentum factor, and the Pastor and Stambaugh (2003) liquidity factor. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

		Q1 (Low)	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Panel A: Full Sample, 2004-2017							
	Alpha	-0.280*	-0.178	-0.127	-0.027	0.145	0.425**
	T stat	[-1.94]	[-1.36]	[-1.00]	[-0.21]	[1.05]	[2.20]
	N	105	100	101	100	97	
	Size	5,440	5,673	5,652	5,961	4,570	
Panel B: Split by Mutual Fund Common Ownership, 2004-2017							
Low	Alpha	-0.506***	-0.224	-0.218	-0.211	-0.021	0.484**
	T stat	[-2.90]	[-1.45]	[-1.36]	[-1.19]	[-0.12]	[1.99]
	N	52	50	50	51	49	
	Size	2,985	2,563	2,694	3,028	2,626	
High	Alpha	-0.171	0.077	0.015	0.172	0.202	0.373
	T stat	[-0.95]	[0.46]	[0.09]	[1.05]	[1.14]	[1.43]
	N	54	49	50	50	47	
	Size	7,532	9,022	8,690	8,818	6,378	
Panel C: Split by Monthly Institutional Abnormal Attention (MIAA) on Customer Firms, 2010-2017							
Low	Alpha	-0.266	-0.179	0.000	-0.005	0.236	0.502*
	T stat	[-1.09]	[-0.68]	[0.00]	[-0.02]	[1.08]	[1.76]
	N	54	52	52	51	50	
	Size	6,259	7,708	6,691	7,493	5,694	
High	Alpha	-0.425*	-0.199	0.087	0.218	-0.116	0.309
	T stat	[-1.80]	[-0.82]	[0.36]	[0.83]	[-0.45]	[0.86]
	N	52	47	46	43	43	
	Size	6,275	5,282	6,395	5,549	4,038	

common ownership group. My result is consistent with the findings in Cohen and Frazzini (2008). The abnormal return in low common ownership group is 0.484% (t-statistic = 1.99) which is statistically significant, and note that the pattern across quintiles is monotonically increasing from the low to the high customer firm return portfolios, but there is no significant abnormal return in L/S trading strategy in high common ownership group.

Panel C of Table 1.3 provides the result of the subsamples with low and high Monthly Institutional investors' Abnormal Attention (MIAA). Here, the abnormal return of customer momentum strategy in low attention group is 0.502% (t-statistic = 1.76), which is marginally statistically significant, whereas the abnormal return in high attention group is 0.309% (t-statistic = 0.86) and statistically insignificant. Note that the sample period is between 2010 and 2017 because the institutional abnormal attention data from Bloomberg is not available until February 2010. This short sample period may explain why there is only marginally significant abnormal returns in low attention group. This result allows me to make a direct comparison between the attention effect from institutional and retail investors, because the construction of monthly institutional and retail investors' abnormal attention is in a similar fashion. In untabulated results, the abnormal return of customer momentum strategy in low and high MRAA group during sample period between 2010 and 2017 is 0.979% (t-statistic=2.93) and 0.145% (t-statistic=0.34) respectively.

In sum, the results in Table 1.2 and 1.3 show that the abnormal return of "customer momentum" trading strategy is larger in the group with low retail and institutional attention. However, the larger difference in abnormal return between retail attention group (0.832% - 0.028%=0.804%) implies that retail attention is more effective in identifying a profitable strategy, and I will turn to a double sorting strategy to drill down into this in Section 1.4.3.

1.4.3 Split Sample Using both Retail and Institutional Investors' Attention

To make an explicit comparison of the effects between retail and institutional attention, I apply a double sort approach to consider both the level of retail and institutional attention. Specifically, I first split the firms into two groups based on the sort of Monthly Retail Abnormal Attention (MRAA), then in each group, I further split the firms based on the sort of mutual fund common ownership or Monthly Institutional Abnormal Attention (MIAA). In this case, I sort the firms into three terciles instead of five quintiles in each subsample due to small sample constraints - dividing into quintiles reduces the number of firms in each category too much for reliable inference. The results are presented in Panel A, B, C and D of Table 1.4.

As a reference point Panel A and C of Table 1.4 provides the anbormal return of the "customer momentum" strategy without sample splitting. Panel A is based on the period between Jun 1st 2004 and May 31st 2017, consistent with the valid sample period of MRAA and mutual fund common ownership. Panel C is based on the period between Jun 1st 2010 and May 31st 2017, consistent with the valid sample period of MIAA. Consistent with previous findings, the customer momentum strategy yields a significant abnormal return. Panel B of Table 1.4 illustrates the abnormal return for the four groups, which correspond to low (high) MRAA with low (high) mutual fund common ownership in the previous month, sorted into portfolios based on low, medium and high customer returns in the previous month. The new double sort analysis shows that when Monthly Retail Abnormal Attention is low, the abnormal returns of customer momentum strategy are 0.788% and 0.500% for low and high mutual fund common ownership respectively, which are much larger than -0.005% and 0.016%, the abnormal return when Monthly Retail Abnormal Attention is high. Panel D of Table 1.4 presents the abnormal return in different level of MRAA and MIAA. The similar pattern shows up as the abnormal return in low MRAA groups

Table 1. 4: Abnormal Return (EW) of Customer Momentum Strategy, Double Sort by Retail and Institutional Investors' Attention

This table provides the abnormal return for calendar-time portfolio for full sample and subsamples double sorted by retail and institutional investors' attention. Panel A and C provides the abnormal return of the portfolios based on the full sample during the period from Jun 1st 2004 to May 31st 2017 and from Jun 1st 2010 to May 31st 2017. Panel B gives the abnormal return based on the subsample double sorted by Monthly Retail Abnormal Attention (MRAA) and Monthly Institutional Abnormal Attention (MIAA). Panel D presents the abnormal return based on the subsample double sorted by Monthly Retail Abnormal Attention (MRAA) and mutual fund common ownership. In each calendar month, supplier firms are sorted based on the return of major customers in the previous month and are grouped into terciles. Low, Medium and High denote the groups corresponding to the lowest, Medium and highest customer return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers.The abnormal return (Alpha) is the intercept on a regression of monthly excess return from the rolling strategy. T stat denotes the t-statistics rounded to two decimals. The explanatory variables are monthly returns of Fama and French (1993) and Fama and French (2015) mimicking portfolios, the Carhart (1997) momentum factor, and the Pastor and Stambaugh (2003) liquidity factor. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

			Low	**Medium**	**High**	**L/S**
Panel A: Full Sample, 2004-2017						
		Alpha	-0.283**	-0.063	0.109	0.392***
		T stat	[-2.31]	[-0.63]	[1.09]	[2.59]
		N	171	165	167	
		Size	5,503	5,857	5,194	
Panel B: Double Sort by MRAA and Common Ownership, 2004-2017						
Low	Low	Alpha	-0.555***	-0.053	0.233	0.788***
		T stat	[-3.06]	[-0.33]	[1.25]	[3.24]
		N	50	48	48	
		Size	2,642	2,692	2,858	
	High	Alpha	-0.094	-0.068	0.406**	0.500**
		T stat	[-0.53]	[-0.38]	[2.37]	[2.10]
		N	52	48	47	
		Size	8,179	8,794	7,591	
High	Low	Alpha	-0.418**	-0.301	-0.423**	-0.005
		T stat	[-2.05]	[-1.49]	[-2.06]	[-0.02]
		N	37	34	34	
		Size	2,758	2,685	2,513	
	High	Alpha	0.011	0.094	0.028	0.016
		T stat	[0.06]	[0.45]	[0.13]	[0.06]
		N	39	34	31	
		Size	8,550	8,440	5,892	

			Low	Medium	High	L/S
Panel C: Full sample, 2010-2017						
		Alpha	-0.285*	-0.067	0.239*	0.524***
		T stat	[-1.79]	[-0.53]	[1.80]	[2.72]
		N	165	162	162	
		Size	6,406	6,752	6,067	
Panel D: Split by MRAA and MIAA, 2010-2017						
Low	Low	Alpha	-0.359	0.077	0.322	0.681*
		T stat	[-1.41]	[0.35]	[1.26]	[1.96]
		Nobs	55	52	51	
		Size	7,425	6,717	6,874	
	High	Alpha	-0.156	0.003	0.439	0.595
		T stat	[-0.64]	[0.01]	[1.60]	[1.56]
		N	49	42	41	
		Size	5,876	5,856	4,924	
High	Low	Alpha	-0.263	0.239	-0.340	-0.076
		T stat	[-1.02]	[0.89]	[-1.03]	[-0.20]
		N	40	35	34	
		Size	7,096	6,075	4,354	
	High	Alpha	-0.251	0.379	-0.573	-0.322
		T stat	[-0.96]	[1.10]	[-1.64]	[-0.83]
		N	36	32	27	
		Size	5,961	6,521	4,139	

are 0.681% and 0.595% with both statistical significance for low and high common owernship group, which are again much larger than -0.076% and -0.322% when MRAA is high. To sum up, Panel B and D of Table 1.4 provide evidence that the effect of retail attention dominate that of institutional attention[5].

To summarize, I find evidence consistent with the price-stablizing role of retail attention, and inconsistent with the price-stablizing role of institutional investor attention. It appears that attention does indeed matter to return persistence and information diffusion, but it is the attention of retail investors that matters. Though the result using mutual fund common ownership suggests that institutional investor attention has an impact, consistent with the findings of Cohen and Frazzini (2008), the double sort analysis by including retail and institutional attention at the same time shows that retail attention absorbs the explanatory power of institutional attention.

1.4.4 Results of Fama Macbeth Forecasting Regression

As a standard tool in asset pricing literature, the Long/Short trading strategy is an intuitive way to show the return predictability between customer and supplier firms. However, it may not be able to draw conclusions on whether the anomaly variable of our interest, the attention of investors, contains unique information while considering other potential explanatory variables. As a robustness test, I use multiple regression in the spirit of Fama and Macbeth (1973) to investigate the marginal effect while controlling for other effects documented in the literature with the following specification,

$$R_t^{sup} = \alpha + \beta * R_{t-1}^{cus} + \gamma' * R_{t-1}^{cus} * \text{High Attention}_{t-1} + \delta' * Z_{t-1} + \varepsilon_t \quad (1.1)$$

[5] The results on double sort approach are similar when I change the order of sorting. Specifically, when I first sort based on Monthly Institutional Abnormal Attention (MIAA) or mutual fund common ownership, then sort on Monthly Retail Abnormal Attention (MRAA), the conclusion does not change.

Where R_t^{sup} is the return of the supplier firm in the current month t; R_{t-1}^{cus} is the lagged return of the customer firm in month t-1; High Attention_{t-1} is a vector of variables capturing the attention using different measures, such as Monthly Retail investors' Abnormal Attention (MRAA), Monthly Institutional investors' Abnormal Attention (MIAA), supplier firm size and mutual fund common ownership; Z_{t-1} is a vector of the control variables. Following Cohen and Frazzini (2008), I include 1-month and 1-year lagged returns of the supplier to control for the one-month reversal effect (Jegadeesh, 1990) and the firms' own momentum effect (Jegadeesh and Titman, 1993). I also include the 1-month lagged return of both the supplier's and the customer's industry portfolio to control for industry momentum (Moskowitz and Grinblatt, 1999) and cross industry momentum effects (Menzly and Ozbas, 2010) respectively. Finally, I also include book to market ratio and the natural logarithm of market capitalization as additional controls (Fama and French, 1993).

Table 1.5 provides the results of the Fama-Macbeth forecasting regression with the specification in equation (1.1). The dependent variable is the return of supplier firms in the current month, and the explanatory variable of primary interest here is the interaction between the return of customer firms and the attention measure in the previous month. The efficient market hypothesis predicts that the coefficient on the lagged customer firm should equal 0, and a delay in information diffusion from customer firms to supplier firms should lead to a positive coefficient on the return of customer firms. If investor attention facilitate information diffusion, the coefficient of the interaction between customer firms' return and attention would be negative. Panel A provides the results based on the full sample period between Jun 1st, 2004 and May 31st, 2017. Column (1) shows the baseline result before introducing the attention variable and confirms the finding in Cohen and Frazzini (2008) that there is a positive and statistically significant coefficient on the

Table 1. 5: Fama Macbeth Forecasting Regression, Comparing MRAA, Common Ownership, and Log(Size)

$$R_t^{sup} = \alpha + \beta * R_{t-1}^{cus} + \gamma_1{}' * R_{t-1}^{cus} * \mathrm{I(High\ MRAA)}_{t-1} + \gamma_2{}' * R_{t-1}^{cus} * \mathrm{I(High\ ComOwn/MIAA)}_{t-1} + \gamma_3{}' * R_{t-1}^{cus} * \mathrm{I(Large\ Size)}_{t-1} + \delta' * Z_{t-1} + \varepsilon_t$$

This table reports the coefficients estimation from Fama Macbeth forecasting regressions as specified in equation (1.1). Panel A provides the results based on the sample period between Jun 1st 2004 and May 31st 2017 and include MRAA, Common Ownership, and firm size as the attention measure. Panel B provides the results based on the sample period between Jun 1st 2010 and May 31st 2017 and include MRAA, MIAA, and firm size as the attention measure. The dependent variable is the supplier firms' return in month t. MRAA_{t-1} denotes the monthly retail investors abnormal attention in month t-1; ComOwn_{t-1} denotes the measure of mutual fund common ownership in month t-1, defined as the number of mutual funds cross hold both customer and supplier firms divided by the total number of mutual funds hold supplier firms; MIAA_{t-1} denotes the monthly institutional investors abnormal attention in month t-1; $\mathrm{Size}_{t-1}$the supplier firm's market capitalization in month t-1. The detail of the coefficients of control variables is provided in Table I.A. 3. T-statistics are calculated based on Newey-West standard errors and are included in parenthesis, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

Panel A: Full Sample (2004-2017)

	(1)	(2)	(3)	(4)	(5)
Ret_{t-1}^{cus}	0.014**	0.031***	0.015*	0.021**	0.032***
	[2.29]	[3.85]	[1.90]	[2.41]	[3.15]
$Ret_{t-1}^{cus} * I(High\ MRAA)_{t-1}$		-0.027***			-0.027***
		[-2.82]			[-2.84]
$Ret_{t-1}^{cus} * I\ (High\ ComOwn)_{t-1}$			0.004		0.006
			[0.41]		[0.54]
$Ret_{t-1}^{cus} * I(Large\ Size)_{t-1}$				-0.007	-0.007
				[-0.65]	[-0.61]
Controls	Yes	Yes	Yes	Yes	Yes
Adjusted R Square	0.067	0.067	0.068	0.068	0.069
Nobs	127133	127133	127133	127133	127133
No. of Firm Pairs	815	815	815	815	815

Panel B: Subsample (2010-2017)

	(1)	(2)	(3)	(4)	(5)
Ret_{t-1}^{cus}	0.023**	0.043***	0.026**	0.027**	0.046***
	[2.57]	[3.73]	[2.36]	[2.15]	[3.06]
$Ret_{t-1}^{cus} * I(High\ MRAA)_{t-1}$		-0.035**			-0.035**
		[-2.58]			[-2.54]
$Ret_{t-1}^{cus} * I(High\ MIAA)_{t-1}$			-0.001		0.003
			[-0.09]		[0.18]
$Ret_{t-1}^{cus} * I(Large\ Size)_{t-1}$				-0.005	-0.006
				[-0.31]	[-0.41]
Controls	Yes	Yes	Yes	Yes	Yes
Adjusted R Square	0.068	0.068	0.069	0.069	0.070
Nobs	68298	68298	68298	68298	68298
No. of Firm Pairs	843	843	843	843	843

lagged return of customer firms. Regarding the control variables, the well documented industry momentum effects ($R_{t-1}^{sup,ind}$), one month return reversal (R_{t-1}^{sup}) and the customer firms' momentum ($R_{t-2,t-12}^{cus}$) all show statistically significant association with the supplier's return. For breivity, I put the coefficients of control variables into Table 1.10, and only present the results of variable of interest in Table 1.5.

Column (2) in panel A of table 1.5 shows the results when an interaction term between the lagged customer return and my core attention measure, Monthly Retail investors' Abnormal Attention (MRAA), is added to the model. If the retail investors' attention affects the return predictability, the higher the attention, the less the return predictability, one should expect a negative and statistically significant loading on the interaction term. The results confirm this conjecture. By using the indicator variable to differentiate the return predictability between the high and low group of monthly retail abnormal attention (MRAA), I find that the group with the low MRAA, the low retail attention group, has a return predictability as high as 3.1% (t-statistic = 3.85), while it drops to 0.4% (3.1% - 2.7% = 0.4%) in the high retail attention group. The results in column (5) confirms that the impact of retail attention on the return predictability is stable while controlling for other competing attention variables such as mutual fund common ownership and supplier's firm size. These findings are consistent with the results I found in the analysis of the "customer momentum" strategy shown in Table 1.2.

Panel B of Table 1.5 shows the results based on a recent subsample between Jun 1st, 2010 and May 31st, 2017, which is the sample period in which the monthly institutional investor's abnormal attention (MIAA) is available. Column (2) and (5) provide results showing that the impact of retail attention on return persistence is stable in this subsample and robust to controlling for institutional attention as proxied by MIAA and firm size. Institutional abnormal attention is again found to be subsumed by retail attention with a coefficient very small in magnitude and statistically insignificant.

1.4.5 Results on Local Retail Investors' Attention

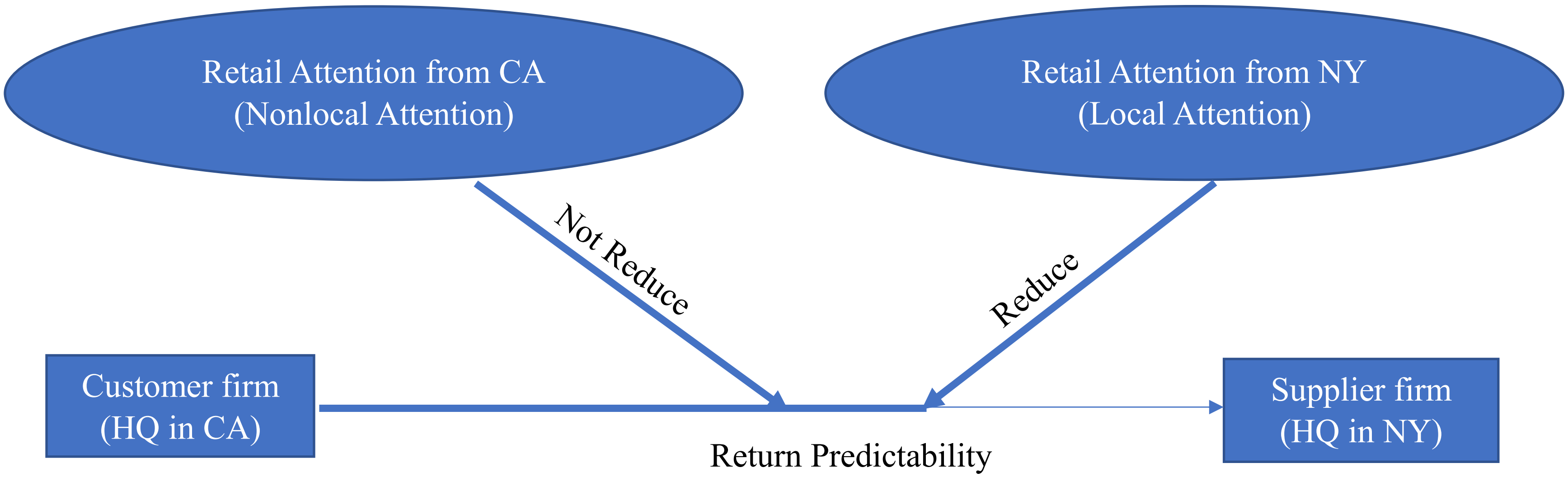
A natural extension of my analysis is to explore heterogeneity of the role of attentions from different types of retail investors. Cziraki, Mondria and Wu (2019) document that local retail investors' attention, defined using geographical information of firms' headquarter, is informed attention. In their conceptual framework, local retail investors will actively collect public information about local firms when they observe private signals. Their empirical results show that asymmetrically high local retail attention relative to national attention predicts returns in the following month. Their results are consistent with recent findings (Kaniel, Saar, and Titman, 2008; Kelley and Tetlock, 2013; Kelley and Tetlock, 2017; Boehmer, Jones, Zhang, and Zhang, 2020) that retail investors could be informed traders. Inspired by these studies, I would like to investigate whether local retail investors' attention is informed attention, and whether their attention facilitates information incorporation.

In identification of the analysis on the impact of local retail attention, I consolidate the structure of customer-supplier network and the "Home Bias" phenomenon (Huberman, 2001; Seasholes and Zhu, 2013; Among others) that local investors overweight local stocks. Figure 1.2 illustrate identification strategy in this test.

To provide an intuitive description of identification strategy, I take an example of a simple customer-supplier relationship with customer firms' headquarter locating in California (CA) and supplier firm's headquarter locating in New York (NY). "Home Bias" phenomenon implies that retail investors located in NY are more likely overweight their investment on the supplier firm in

Figure 1. 2: Return Predictability and Local Attention

This Figure provides an example to illustrate the relationship between return predictability and local attention. The return predictability is from customer firm with headquarter locating in California (CA) to supplier firm with headquarter locating in New York (NY). Retail attention from New York (NY) is identified as local retail attention because it is from the same state as supplier firm headquarter locates. Retail attention from California (CA) is identified as one of nonlocal retail attention because it is not from the same state as supplier firm headquarter locates. Technically, all the states other than NY in this case are nonlocal retail attention.



this example. In another words, abnormally high attention paid on customer firms from NY is more likey paid by supplier firms' local investors. So, I use retail attention from NY as a proxy of local retail investor's attention on supplier firms, and explore how return predictability varies with the level of this local attention. If local retail attention is informed, high attention on customer firms by supplier firms' local retail investors will facilitate information diffusion between customer and supplier firms and reduce the return predictability in this context. Here, I also choose retail attention from CA as nonlocal retail attention and explore how return predictability varies with the level of nonlocal attention. Leaning on Cziraki, Mondria and Wu (2019), this is a good candidate for a placebo test as the attention from CA is more likely paid by customer firms' local investors, whose attention should be uninformed with respect to the supplier firms, and thus have no impact on the information diffusion between customer and supplier firms.

The results with regards to the impact of local and nonlocal retail attention on the return predictability are provided in Table 1.6. Panel B of Table 1.6 shows how the return predictability are different between high and low local retail attention group. Specifically, there is a monotonic pattern on the abnormal return from Q1 to Q5 in the group with low local retail attention. The zero-cost L/S trading strategy yield both economically and statistically significant abnormal return and both long and short leg shows significant abnormal return. In contrast, there is no statistically significant profit in L/S trading strategy when local retail attention is high. These results are consistent with my conjecture that when local retail investors attention pay attention to customer firms, the information transmission will be fast and leaves little return persistence.

As a placebo test, Panel C of Table 1.6 provides the results with the variation of nonlocal retail attention, proxied by the retail attention from the customer firms' headquarter location. "Home Bias" literature suggests that this nonlocal retail attention is more likely paid by customer

Table 1. 6: Abnormal Return (EW) of Customer Momentum Strategy, Single Sort by Local and Nonlocal Retail Investors Attention

This table provides the equal weighted abnormal return of customer momentum strategy for calendar-time portfolio for full sample and subsamples splitted by the level of local and nonlocal retail investors' attention. Panel A provides the abnormal return of the portfolios based on the full sample. Panel B and C presents abnormal return in the groups of high and low local and nonlocal retail attention. Local (Nonlocal) retail attention is defined as the attention paid on customer firms by retail investors reside in the same states that supplier (customer) firms' headquater locates. In each calendar month, supplier firms are sorted based on the return of major customers in the previous month and are grouped into five quintiles. Q1 and Q5 denote the groups corresponding to the lowest and highest customer return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers.The abnormal return (Alpha) is the intercept on a regression of monthly excess return from the rolling strategy. T stat denotes the t-statistics rounded to two decimals. The explanatory variables are monthly returns of Fama and French (1993) and Fama and French (2015) mimicking portfolios, the Carhart (1997) momentum factor, and the Pastor and Stambaugh (2003) liquidity factor. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

		Q1 (Low)	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Panel A: Full Sample, 2004-2017							
	Alpha	-0.280*	-0.178	-0.127	-0.027	0.145	0.425**
	T stat	[-1.94]	[-1.36]	[-1.00]	[-0.21]	[1.05]	[2.20]
	Nobs	105	100	101	100	97	
	Size	5,440	5,673	5,652	5,961	4,570	
Panel B: Split Sample by Supplier Firms' Local Retail Attention on Customer Firms, 2004-2017							
Low	Coeff	-0.293*	-0.167	-0.057	0.037	0.323**	0.617***
	Tstat	[-1.73]	[-1.15]	[-0.37]	[0.24]	[2.12]	[2.93]
	Nobs	63	59	59	59	58	
	Size	5,473	5,121	5,516	5,589	4,294	
High	Coeff	-0.427**	-0.143	-0.189	0.097	-0.138	0.289
	Tstat	[-2.20]	[-0.81]	[-1.10]	[0.56]	[-0.64]	[1.00]
	Nobs	44	43	42	42	40	
	Size	5,616	6,431	5,837	6,425	4,947	
Panel C: Split Sample by Supplier Firms' Nonlocal Retail Attention on Customer Firms, 2004-2017							
Low	Coeff	-0.268	-0.240	0.024	0.068	-0.038	0.229
	Tstat	[-1.44]	[-1.40]	[0.14]	[0.35]	[-0.22]	[0.93]
	Nobs	61	56	55	56	54	
	Size	4,962	5,362	5,099	5,485	4,164	
High	Coeff	-0.356*	-0.005	-0.423**	0.161	0.000	0.356
	Tstat	[-1.92]	[-0.03]	[-2.32]	[0.80]	[0.00]	[1.37]
	Nobs	49	45	47	45	42	
	Size	5,456	6,025	5,767	5,742	4,250	

firms' local retail investors, not supplier firms' local retail investors. So this nonlocal retail attention could help facilitate customer firms' information into customer stock prices but should not directly help facilitate customer firms' news into supplier stock prices. Thus, varying the level of this nonlocal retail attention should not have direct impact on the return persistence between customer and supplier stocks. The results are consistent with this conjecture. The results in Panel C of Table 1.6 reveal that there is no statistically significant abnormal return in either the high or low nonlocal attention portfolios.

Finally, I examine the impact of local retail attention on return predictability using Fama Macbeth forecasting regression. Table 1.7 report the results with a single interaction term between lagged customer stocks' return and a broad range of relevant attention measures, The core variable of interest is local retail attention. A series of meaningful attention measures serve as control variables are nonlocal retail attention and supplier firms' size. The nonlocal retail attention measure is an ideal candidate for a placebo test showing the impacts from another regional retail attention. For breivity, I put the coefficients of control variables into Table 1.11, and only present the results of variable of interest in Table 1.7.

Column 1 of Table 1.7 reports that "customer momentum" return predictability weakens when local retail attention is high. The return predictability drops from 2.6% to 0.6% (2.6%-2.0%=0.6%), with statistical insignificance. Compared to column 1, column 2 shows that this return predictability does not have statistically significant difference when nonlocal retail attention is high. Finally, Column 4 present results by incorporating all of these attention measures into Fama Macbeth regression and confirms that the findings are robust and remain unchanged.

Table 1. 7: Fama Macbeth Forecasting Regression, Comparing Local and Nonlocal Abnormal Attention

$$R_t^{sup} = \alpha + \beta * R_{t-1}^{cus} + \gamma_1' * R_{t-1}^{cus} * \mathrm{I(High\ Local\ Abatten)}_{t-1} + \gamma_2' * R_{t-1}^{cus} * \mathrm{I(High\ Nonloacal\ Abatten)}_{t-1} + \gamma_3{}' * R_{t-1}^{cus} * \mathrm{I(Large\ Size)}_{t-1} + \delta' * Z_{t-1} + \varepsilon_t$$

This table reports the coefficients estimation from Fama Macbeth forecasting regressions as specified in equation (1.1) in the sample period between Jun 1st 2004 and May 31st 2017. The dependent variable is the supplier firms' return in month t. All the attention measures are indicator variables equal one if the value is above median level. Local (nonlocal) abatten represent the abnormal attention paid on customer firms by retail investors reside in the same states that supplier (customer) firms' headquater locates. Abnormal attention is calculated as the natural logorithms difference between SVI in current month and the median SVI during previous quarter. The detail of the coefficients of control variables is provided in Table I.A. 4. T-statistics are calculated based on Newey-West standard errors and are included in parenthesis, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

	(1)	(2)	(3)	(4)
Ret_{t-1}^{cus}	0.026***	0.024***	0.021**	0.033***
	[3.66]	[2.99]	[2.41]	[3.42]
$Ret_{t-1}^{cus} * I(High\ Local\ Abatten)_{t-1}$	-0.020**			-0.020**
	[-2.53]			[-2.41]
$Ret_{t-1}^{cus} * I(High\ Nonlocal\ Abatten)_{t-1}$		-0.016		-0.006
		[-1.57]		[0.56]
$Ret_{t-1}^{cus} * I(Large\ Size)_{t-1}$			-0.007	-0.008
			[-0.65]	[-0.78]
Controls	Yes	Yes	Yes	Yes
Adjusted R Square	0.067	0.067	0.068	0.068
Nobs	127133	127133	127133	127133
No of Firm-pairs	815	815	815	815

To summarize, I find that local retail attention plays a positive role on information efficiency in the information transmission from customer to supplier firms. A placebo test shows nonlocal retail attention, which is expected to be uniformed, has no impact on return persistence.

1.5 Robustness

1.5.1 Results after correcting for the biased alpha between economically linked firms

I verify whether the retail investor attention results in my study are immune to the recent criticism on the "customer momentum" strategy. Burt and Hrdlicka (2021) argue that if customers and suppliers have correlated risk exposure, which seems likely, then sorting suppliers by their customers' past excess returns produces a long/short portfolio which mechanically has a positive, persistent return. In other words, to the extent that the suppliers expected excess returns are mismeasured, perhaps due to the misspecified asset pricing model, then this long/short portfolio does not simply have a persistent positive return but also a persistent positive alpha - "evidence" of information persistence. Then the observed positive alpha may not reflect a return persistence from the customer firms.

To fix this issue, Burt and Hrdlicka (2021) and Lee, Sun, Wang, and Zhang (2019) suggest using the idiosyncratic return to replace the raw return in the sorting stage. Specifically, they use the daily returns of customer firms from the previous 12 months and Fama French four-factor model (Fama and French, 1993; Carhart, 1997) to estimate the factor loadings, and then use the loading and actual returns to calculate the idiosyncratic return in month t-1. Lastly, they use the obtained idiosyncratic return, which should screen out the potential correlation in alpha, to make the sort.

The results in Panel A and B of Table 1.8 are qualitatively identical with those in Panel A and B of Tables 2. Specifically, Panel A shows that the customer momentum strategy yields 0.423% monthly abnormal return in the full sample. Results presented in Panel B confirm the findings in Panel B of Table 1.2 that the group with low retail attention has the largest abnormal return, which is statistically significantly different from those in the group with high retail attention.

1.5.2 Split Sample Using both Firm Size and Attention

One concern is that the findings on monthly retail investors' abnormal attention may also be driven by other factors, such as size effects. Although I have controlled for firm size in the Fama Macbeth regressions reported in Tables 1.5, I also make use of a double sort approach similar to the one used in section 1.4.3 to address this concern.

I conduct double sort analysis by first sorting supplier firms into two groups based on firm size, proxied as firms' market capitalization. Then, in each subgroup, I sort the supplier firms based on the MRAA. By doing so, I end up with four groups of suppliers and implement the "customer momentum" strategy in each group. The results are presented in Table 1.9.

Panel A of Table 1.9 provides the abnormal return of the "customer momentum" strategy which is identical to Panel A of Table 1.4 for the ease of comparison. Panel B of Table 1.5 illustrates the abnormal return for the four groups, which correspond to small (large) supplier firms with low (high) retail attention paid to the customer firm in the previous month. The new double sort analysis shows that both firm size and the attention of retail investors are important sources of return predictability. In the small firm group, the low attention subgroup exhibits the strongest

Table 1. 8: Abnormal Return (EW) of Customer Momentum Strategy, Sort by Retail Investors' Attention

This table provides the equal weighted abnormal return of customer momentum strategy for calendar-time portfolio for full sample and subsamples splitted by the level of Monthly Retail Abnormal Attention (MRAA) on customer firms. Panel A provides the abnormal return of the portfolios based on the full sample. Panel B shows the abnormal return in the groups of high and low Monthly Retail Abnormal Attention (MRAA) on customer firms. In each calendar month, supplier firms are sorted based on the idiosyncratic return of major customers in the previous month and are grouped into five quintiles. The use of idiosyncratic return is suggested by Burt and Hrdlicka (2016). Q1 and Q5 denote the groups corresponding to the lowest and highest customer idiosyncratic return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers. The abnormal return (Alpha) is the intercept from a multiple time series regression including risk factors documented in Fama and French (1993), Carhart (1997), Paster and Stambaugh (2003), and Fama and French (2015), T stat denotes the t-statistics rounded to two decimals. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

		Q1 (Low)	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Panel A: Full Sample, 2004-2017, Idiosyncratic Return Approach							
	Alpha	-0.234*	-0.166	-0.174	-0.034	0.189	0.423**
	T stat	[-1.70]	[-1.21]	[-1.41]	[-0.28]	[1.42]	[2.33]
	N	106	100	100	100	97	
	Size	5,290	5,717	5,635	6,063	4,643	
Panel B: Split by Monthly Retail Abnormal Attention (MRAA) on Customer Firms, 2004-2017, Idiosyncratic Return Approach							
Low	Alpha	-0.315*	-0.042	-0.264	0.065	0.394**	0.708***
	T stat	[-1.79]	[-0.27]	[-1.45]	[0.40]	[2.14]	[2.85]
	N	63	60	57	58	56	
	Size	5,265	5,347	4,963	6,639	4,688	
High	Alpha	-0.158	-0.088	-0.221	-0.029	-0.156	0.002
	T stat	[-0.92]	[-0.44]	[-1.04]	[-0.13]	[-0.74]	[0.01]
	N	47	43	42	42	38	
	Size	4,940	5,560	5,691	4,857	3,712	

Table 1. 9: Abnormal Return (EW) of Customer Momentum Strategy, Double Sort by Firm Size and Retail Investors' Attention

This table provides the abnormal return for calendar-time portfolio for full sample and subsamples double sorted by firm size and Retail investors' attention. Panel A provides the abnormal return of the portfolios based on the full sample during the period from Jun 2004 to May 2017. Panel B gives the abnormal return based on the subsample double sorted by supplier firms' size (market capitalization) and Monthly Retail Abnormal Attention (MRAA). In each calendar month, supplier firms are sorted based on the return of major customers in the previous month and are grouped into terciles. Low, Medium and High denote the groups corresponding to the lowest, Medium and highest customer return. L/S is the zero-cost portfolio that buy the supplier stocks linked with the best performed customers and sell short the supplier stocks linked with the worst performed customers.The abnormal return (Alpha) is the intercept on a regression of monthly excess return from the rolling strategy. T stat denotes the t-statistics rounded to two decimals. The explanatory variables are monthly returns of Fama and French (1993) and Fama and French (2015) mimicking portfolios, the Carhart (1997) momentum factor, and the Pastor and Stambaugh (2003) liquidity factor. N provides the monthly average number of supplier firms. Size provides the monthly average supplier firms' market capitalization in previous month in each portfolio, with the units of million of US dollars. EW denotes that the portfolio return is equal weighted. The abnormal returns are in monthly percent, the t-statistics are included in parentheses, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

			Low	**Medium**	**High**	**L/S**
Panel A: Full sample, 2004-2017						
		Alpha	-0.283**	-0.063	0.109	0.392***
		T stat	[-2.31]	[-0.63]	[1.09]	[2.59]
		N	171	165	167	
		Size	5,503	5,857	5,194	
Panel B: Double Sort by Firm Size and MRAA, 2004-2017						
Small	Low	Alpha	-0.540***	-0.144	0.560***	1.100***
		T stat	[-2.68]	[-0.85]	[2.85]	[4.27]
		N	49	48	48	
		Size	379	383	381	
	High	Alpha	-0.395*	-0.147	-0.321	0.074
		T stat	[-1.91]	[-0.73]	[-1.55]	[0.25]
		N	37	35	34	
		Size	392	402	398	
Large	Low	Alpha	-0.086	-0.240	0.206	0.291
		T stat	[-0.46]	[-1.55]	[1.26]	[1.19]
		N	53	48	47	
		Size	10,560	11,137	10,374	
	High	Alpha	-0.155	0.121	-0.047	0.108
		T stat	[-0.93]	[0.62]	[-0.22]	[0.39]
		N	38	33	31	
		Size	10,833	9,827	8,332	

long/short return of 1.100%. Small size is not, however, enough. Small firms that were associated with customer firms receiving high retail attention last month show an insignificant long/short abnormal return of 0.074%, suggesting that when retail investors pay high attention to the customer firms, the small suppliers do not exhibit return predictability. As a comparison, in the large firm group, neither the low nor high attention subgroup show a significant abnormal return. In summary, the main finding based on this double sort analysis is that both the firm size and retail investors attention matters in the case of return predictability. In other words, the main effect of return predictability is concentrated on smaller firms with low attention from retail investors, which is consistent with the effect's strength based on the single sort method when using equal weighted portfolios versus value-weighted portfolios, discussed above.

1.6 Discussion

It is worthwhile to discuss why this return predictability is not eliminated by trading from institutional investors, even after being documented by Cohen and Frazzini (2008), among others. There are several possible reasons for this finding. At first, institutional investors may choose to neglect smaller firms, purely for the consideration of the optimal allocation of attention. Chuprinin, Gorbenko, and Kang (2019) propose a theoretical model explaining why investors with more assets under management will rationally ignore smaller firms. Specifically, under an assumption of mispricing equilibrium, all firms are expected to have the same level of expected dollar arbitrage profit. Institutional investors prefer not to spend investigative resources on the smaller firm because the dollar revenue of arbitrage profit is lower for a smaller firm, given the same level of mispricing.

Secondly, it is possible that most of the institutional investors are unable to engage in the trading of these supplier firms due to investment policy or preference. Even if an institutional investor faced no trading restrictions on these stocks, any meaningful trade would likely trigger the reporting requirement of SEC. Take the most profitable subsample in panel B Table 1.9 as an example. Table 1.9 reports that the group with small suppliers and low attention of retail investors shows an abnormal return of 1.100% per month. The average market capitalization, in this case, is roughly 380 million US dollars. There are around 50 firms involved in the long/short trading in the bottom and top tercile groups (data on the number of firms in each portfolio is not displayed in the table, but available on request). Provided that 5% is the maximum ownership stake that does not trigger the reporting procedure[6], and further assuming that a purchase of this magnitude will not significantly change the price (a very aggressive assumption exaggerating the profitability of this strategy), then the abnormal profit by exercising the customer momentum strategy will be less than 9 million (380 million x 5% x 50 firms x 1.100% ≈ 10.45 million) in any given month. Given the large amount of funds under the management of institutional investors, this amount of profit, were it even feasible given transactions costs, is unlikely to motivate institutional investors to trade actively on this mispricing.

1.7 Conclusion

This paper finds that the limited attention of retail investors is the main source resulting in return persistence in the context of customer and supplier firms, a group of firms that constitute roughly 58.4% of all publically traded equities in the U.S. I show that the attention of retail

[6] SEC require shareholders who acquire more than 5% of the outstanding shares of that class must file beneficial owner reports on Schedule 13D or 13G until their holdings drop below 5%.

investors matters, that with this attention we see market efficiency, and in absence of this attention we see mispricing – return persistence. In addition I find that the institutional investors' attention does not appear to be matter if we control for retail investor attention, with these results driven by small firms, a group of firms that are unlikely to attract the attention of institutional investors, likely due to limited profits from trading activity on these firms.

The most important implication of this study is for the price-stabilizing impact of retail investors' attention. Retail investors are broadly recognized as the traders that are most subject to severe behavioral biases, such as price overreaction or net buying imbalance on attention-grabbing events, or noise traders whose trading does not help improve market efficiency. However, my findings in this paper show that retail investor activity can facilitate information incorporation to stock prices, which in turn, improves market efficiency. In particular, I find that local retail attention, which is documented as informed attention in previous studies, has price-stablizing impacts on financial markets. This implies that some retail investors are smart investors and can stabilize financial markets. On the contrary, institutional investors and their attention is not material in this setting. Examining controls for firm size shows that the impact of retail attention is mainly apparent for smaller firms. In other words, retail attention appears to complement institutional attention by moving the market in the right direction for firms which are most likely ignored by institutional investors.

Table 1. 10: Fama Macbeth Forecasting Regression

$$R_t^{sup} = \alpha + \beta * R_{t-1}^{cus} + \gamma_1{}' * R_{t-1}^{cus} * \mathrm{I(High\ MRAA)}_{t-1} + \gamma_2{}' * R_{t-1}^{cus} * \mathrm{I(High\ ComOwn/MIAA)}_{t-1} + \gamma_3{}' * R_{t-1}^{cus} * \mathrm{I(Large\ Size)}_{t-1} + \delta' * Z_{t-1} + \varepsilon_t$$

This table reports the coefficients estimation from Fama Macbeth forecasting regressions as specified in equation (1.1). Panel A provides the results based on the sample period between Jun 2004 and May 2017 and include MRAA, Common Ownership, and firm size as the attention measure. Panel B provides the results based on the sample period between Jun 2010 and May 2017 and include MRAA, MIAA, and firm size as the attention measure. The dependent variable is the supplier firms' return in month t. MRAA_{t-1} denotes the monthly retail investors abnormal attention in month t-1; MIAA_{t-1} denotes the monthly institutional investors abnormal attention in month t-1; ComOwn_{t-1} denotes the measure of mutual fund common ownership in month t-1, defined as the number of mutual funds cross hold both customer and supplier firms divided by the total number of mutual funds hold supplier firms; $\mathrm{Size}_{t-1}$the supplier firm's market capitalization in month t-1. Column 2-5 shows the results using indicator variables which equal one if the value is above the median level. Column 6-9 show the results using continuous value of attention variables. R_{t-1}^{cus} and R_{t-1}^{sup} denotes the customer and supplier firms' return in month t-1; $R_{t-2,t-12}^{cus}$ and $R_{t-2,t-12}^{sup}$ denotes customer and supplier firms' one year momentum return skipped the month t-1; $R_{t-1}^{cus,ind}$ and $R_{t-1}^{sup,ind}$ denotes customer and supplier firms' industry return in month t-1; $R_{t-2,t-12}^{cus,ind}$ and $R_{t-2,t-12}^{sup,ind}$ denotes customer and supplier firms' one year momentum return skipped the month t-1, along with the log form of supplier firms' market capitalization and book to market ratio. T-statistics are calculated based on Newey-West standard errors and are included in parenthesis, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

Panel A: Full sample (2004-2017)		Indicator Variable				Continuous Variable			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
R_{t-1}^{cus}	0.014**	0.031***	0.015*	0.021**	0.032***	0.039***	0.014	0.058**	0.079***
	[2.29]	[3.85]	[1.90]	[2.41]	[3.15]	[4.26]	[1.49]	[2.20]	[3.01]
$R_{t-1}^{cus} * \text{MRAA}_{t-1}$		-0.027***			-0.027***	-0.021***			-0.021***
		[-2.82]			[-2.84]	[-3.42]			[-3.44]
$R_{t-1}^{cus} * \text{ComOwn}_{t-1}$			0.004		0.006		0.004		0.004
			[0.41]		[0.54]		[0.54]		[0.59]
$R_{t-1}^{cus} * \text{Size}_{t-1}$				-0.007	-0.007			-0.043*	-0.045*
				[-0.65]	[-0.61]			[-1.66]	[-1.75]
$R_{t-2,t-12}^{cus}$	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.006***
	[2.85]	[2.96]	[2.96]	[3.01]	[3.08]	[3.01]	[2.98]	[3.00]	[3.08]
R_{t-1}^{sup}	-0.016*	-0.015*	-0.015*	-0.015*	-0.015*	-0.016*	-0.015*	-0.015*	-0.015*
	[-1.94]	[-1.79]	[-1.75]	[-1.75]	[-1.71]	[-1.83]	[-1.76]	[-1.74]	[-1.74]
$R_{t-2,t-12}^{sup}$	0.000	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
	[-0.04]	[-1.12]	[-1.12]	[-1.10]	[-1.09]	[-1.12]	[-1.11]	[-1.12]	[-1.10]
$R_{t-1}^{cus,ind}$	0.050*	0.033	0.039	0.039	0.036	0.035	0.039	0.037	0.035
	[1.78]	[1.06]	[1.23]	[1.26]	[1.16]	[1.10]	[1.23]	[1.19]	[1.11]
$R_{t-2,t-12}^{cus,ind}$	0.000	0.000	0.001	0.001	0.001	0.000	0.001	0.001	0.001
	[0.01]	[0.04]	[0.11]	[0.12]	[0.14]	[0.01]	[0.12]	[0.12]	[0.12]
$R_{t-1}^{sup,ind}$	0.082**	0.094**	0.093**	0.093**	0.094**	0.093**	0.094**	0.095**	0.095**
	[2.34]	[2.50]	[2.48]	[2.47]	[2.50]	[2.48]	[2.49]	[2.49]	[2.52]
$R_{t-2,t-12}^{sup,ind}$	0.003	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002
	[0.31]	[0.21]	[0.21]	[0.22]	[0.19]	[0.24]	[0.22]	[0.22]	[0.23]
Log(Size)_{t-1}	-0.019	-0.024	-0.027	-0.040	-0.041	-0.024	-0.029	-0.034	-0.035
	[-0.41]	[-0.50]	[-0.56]	[-0.79]	[-0.83]	[-0.48]	[-0.58]	[-0.66]	[-0.69]
B/M	-0.024	-0.163	-0.166	-0.160	-0.168	-0.159	-0.170	-0.160	-0.169
	[-0.12]	[-0.79]	[-0.81]	[-0.78]	[-0.82]	[-0.77]	[-0.82]	[-0.78]	[-0.82]
Adjusted R Square	0.067	0.067	0.068	0.068	0.069	0.067	0.068	0.068	0.069
Nobs	127133	127133	127133	127133	127133	127133	127133	127133	127133
No of Firm-pairs	815	815	815	815	815	815	815	815	815

Panel B: Subsample (2010-2017)		Indicator Variable				Continuous Variable			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
R_{t-1}^{cus}	0.023**	0.043***	0.026**	0.027**	0.046***	0.043***	0.027**	0.077**	0.096***
	[2.57]	[3.73]	[2.36]	[2.15]	[3.06]	[3.24]	[2.00]	[2.22]	[2.62]
$R_{t-1}^{cus} * \text{MRAA}_{t-1}$		-0.035**			-0.035**	-0.016**			-0.017**
		[-2.58]			[-2.54]	[-1.99]			[-2.13]
$R_{t-1}^{cus} * \text{MIAA}_{t-1}$			-0.001		0.003		0.002		0.007
			[-0.09]		[0.18]		[0.17]		[0.70]
$R_{t-1}^{cus} * \text{Size}_{t-1}$				-0.005	-0.006			-0.053*	-0.054*
				[-0.31]	[-0.41]			[-1.71]	[-1.75]
$R_{t-2,t-12}^{cus}$	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001
	[0.69]	[0.59]	[0.52]	[0.60]	[0.52]	[0.60]	[0.71]	[0.53]	[0.66]
R_{t-1}^{sup}	-0.008	-0.009	-0.009	-0.009	-0.008	-0.009	-0.009	-0.009	-0.009
	[-0.69]	[-0.76]	[-0.75]	[-0.73]	[-0.70]	[-0.80]	[-0.78]	[-0.75]	[-0.77]
$R_{t-2,t-12}^{sup}$	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	[0.9]	[0.49]	[0.49]	[0.51]	[0.53]	[0.49]	[0.50]	[0.51]	[0.53]
$R_{t-1}^{cus,ind}$	0.094**	0.068*	0.072*	0.077*	0.071*	0.071*	0.070*	0.074*	0.068
	[2.46]	[1.65]	[1.74]	[1.85]	[1.71]	[1.71]	[1.70]	[1.77]	[1.64]
$R_{t-2,t-12}^{cus,ind}$	0.002	0.005	0.005	0.005	0.006	0.005	0.004	0.005	0.004
	[0.25]	[0.54]	[0.53]	[0.48]	[0.61]	[0.51]	[0.37]	[0.48]	[0.43]
$R_{t-1}^{sup,ind}$	0.079	0.075	0.074	0.074	0.073	0.075	0.072	0.076	0.073
	[1.46]	[1.31]	[1.29]	[1.28]	[1.28]	[1.31]	[1.25]	[1.31]	[1.27]
$R_{t-2,t-12}^{sup,ind}$	0.024**	0.031**	0.030**	0.031**	0.031**	0.031**	0.030**	0.030**	0.030**
	[1.99]	[2.33]	[2.32]	[2.35]	[2.33]	[2.33]	[2.29]	[2.30]	[2.28]
Log(Size)_{t-1}	-0.016	-0.006	-0.008	-0.025	-0.028	-0.006	-0.008	-0.008	-0.012
	[-0.25]	[-0.10]	[-0.13]	[-0.39]	[-0.44]	[-0.10]	[-0.12]	[-0.12]	[-0.18]
B/M	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	[0.66]	[0.55]	[0.53]	[0.61]	[0.54]	[0.55]	[0.55]	[0.60]	[0.53]
Adjusted R Square	0.068	0.068	0.069	0.069	0.070	0.068	0.068	0.068	0.069
Nobs	68298	68298	68298	68298	68298	68298	68298	68298	68298
No of Firm-pairs	843	843	843	843	843	843	843	843	843

Table 1. 11: Fama Macbeth Forecasting Regression, Comparing Local and Nonlocal Abnormal Attention

$$R_t^{sup} = \alpha + \beta * R_{t-1}^{cus} + \gamma_1' * R_{t-1}^{cus} * \text{I(High Local Abatten)}_{t-1} + \gamma_2' * R_{t-1}^{cus} * \text{I(High Nonloacal Abatten)}_{t-1} + \gamma_3{}' * R_{t-1}^{cus} * \text{I(Large Size)}_{t-1} + \delta' * Z_{t-1} + \varepsilon_t$$

This table reports the coefficients estimation from Fama Macbeth forecasting regressions as specified in equation (1.1) in the sample period between Jun 1st 2004 and May 31st 2017. The dependent variable is the supplier firms' return in month t. All the attention measures are indicator variables equal one if the value is above median level. Local (nonlocal) abnormal attention represent the abnormal attention paid on customer firms by retail investors reside in the same states that supplier (customer) firms' headquater locates. National abnormal attention represent the abnormal attention paid on customer firms by U.S. national wide retail investors. Abnormal attention is calculated as the natural logorithms difference between SVI in current month and the median SVI during previous quarter. R_{t-1}^{cus} and R_{t-1}^{sup} denotes the customer and supplier firms' return in month t-1; $R_{t-2,t-12}^{cus}$ and $R_{t-2,t-12}^{sup}$ denotes customer and supplier firms' one year momentum return skipped the month t-1; $R_{t-1}^{cus,ind}$ and $R_{t-1}^{sup,ind}$ denotes customer and supplier firms' industry return in month t-1; $R_{t-2,t-12}^{cus,ind}$ and $R_{t-2,t-12}^{sup,ind}$ denotes customer and supplier firms' one year momentum return skipped the month t-1, along with the log form of supplier firms' market capitalization and book to market ratio. T-statistics are calculated based on Newey-West standard errors and are included in parenthesis, and 10%, 5% and 1% statistical significance is indicated by *, ** and *** respectively.

	(1)	(2)	(3)	(4)
Ret_{t-1}^{cus}	0.026***	0.024***	0.021**	0.033***
	[3.66]	[2.99]	[2.41]	[3.42]
$Ret_{t-1}^{cus} * I(High\ Local\ Abatten)_{t-1}$	-0.020**			-0.020**
	[-2.53]			[-2.41]
$Ret_{t-1}^{cus} * I(High\ Nonlocal\ Abatten)_{t-1}$		-0.016		-0.006
		[-1.57]		[-0.56]
$Ret_{t-1}^{cus} * I(Large\ Size)_{t-1}$			-0.007	-0.008
			[-0.65]	[-0.78]
$R_{t-2,t-12}^{cus}$	0.005***	0.006***	0.005***	0.006***
	[2.95]	[3.06]	[3.01]	[3.09]
R_{t-1}^{sup}	-0.015*	-0.015*	-0.015*	-0.015*
	[-1.79]	[-1.77]	[-1.75]	[-1.74]
$R_{t-2,t-12}^{sup}$	-0.003	-0.003	-0.003	-0.003
	[-1.13]	[-1.12]	[-1.10]	[-1.10]
$R_{t-1}^{cus,ind}$	0.038	0.037	0.039	0.038
	[1.20]	[1.19]	[1.26]	[1.20]
$R_{t-2,t-12}^{cus,ind}$	0.000	0.001	0.001	0.001
	[0.06]	[0.07]	[0.12]	[0.15]
$R_{t-1}^{sup,ind}$	0.094**	0.093**	0.093**	0.094**
	[2.48]	[2.47]	[2.47]	[2.48]
$R_{t-2,t-12}^{sup,ind}$	0.002	0.002	0.002	0.001
	[0.20]	[0.17]	[0.22]	[0.14]
Log(Size)$_{t-1}$	-0.024	-0.025	-0.040	-0.039
	[-0.50]	[-0.50]	[-0.79]	[-0.78]
B/M	-0.155	-0.163	-0.160	-0.157
	[-0.75]	[-0.79]	[-0.78]	[-0.77]
Adjusted R Square	0.067	0.067	0.068	0.068
Nobs	127133	127133	127133	127133
No of Firm-pairs	815	815	815	815

Chapter 2 Does Social Interaction Spread Fear among Institutional Investors? Evidence from COVID-19

2.1 Introduction

Against the backdrop of the rapidly growing field of social economics and finance (Hirshleifer 2020), recent work has documented the impact of investor social connectivity on portfolio selections. Social networks transmit both value-relevant information and cognitive biases between investors and firm managers, thereby affecting investor behavior and portfolio decisions. There is evidence that institutional investors acquire an investment edge from interacting with corporate executives or board members through alumni networks (Cohen et al. 2008; Hong and Xu 2019). However, there is also evidence that fund managers who are socially connected to the firms they invest in do not earn superior returns (Kuchler et al. 2020). Furthermore, social interactions aggravate behavioral biases for retail investors with respect to lottery stocks (Bali et al. 2019).

In this paper, we take a new look at the impact of social connectedness on active all-equity mutual fund manager behavior during the COVID-19 pandemic outbreak. We identify COVID-19 "hotspot" counties in the US, as well as counties that are highly socially connected to these hotspots, during the first quarter of 2020, and we ask whether social connectedness is associated with informed or panic-driven trading behaviors. Under what we call the *smart connection hypothesis*, being socially connected to COVID-19 hotspots allows fund managers a pathway to gain valuable insights about the pandemic and act accordingly, whereas under the *salience hypothesis*, social connection to the highly salient outbreak causes these same managers to focus on the negative outcomes of COVID-19 and in turn, make suboptimal trading decisions. We distinguish the two hypotheses by examining fund performance in the quarter subsequent to the initial viral outbreak in the first quarter.

Several features of the COVID-19 episode makes it a unique window for our study. First, the pandemic outbreak caused unprecedented fear and movements in the financial markets, which creates a rare opportunity to examine intensified investor biases such as salience bias associated with social interactions. Second, this event was a sudden exogenous shock to the economy, and the relationship between social connectiveness to COVID hotspots and stock selling during the event should not be due to endogeneity. Third, granular data on COVID-19 cases and Facebook social connectedness are available to identify hotspot counties and counties with high and low connectedness to the hotpots. This provides a sizable variation of the connectedness score across counties, so that there is a control in the experiment to address endogeneity. Finally, the dramatic selloff and rebound of the stocks during our event period allows a quick resolution of the two competing hypotheses.

We identify counties with at least 2,000 cumulative cases as of March 30, 2020 as COVID-19 hotspot counties at the end of the first quarter of 2020. A total of 7 counties are identified as hotspots, covering the vicinities of New York City, Chicago, Los Angeles, and Seattle. We then use Facebook social connectedness index (SCI) to further classify non-hotspot counties into the high- and low-SCI groups based on their social connections to hotspot counties.

Our empirical investigations consist of two parts. In the first part, we examine whether being located in or socially connected to COVID hotspots leads to heavy stock selling by fund managers. We find that during the COVID outbreak quarter of 2020Q1, both being in the hotspot itself and being socially connected to these hotspots intensified institutional stock selling. Multivariate tests indicate that fund managers in the hotspot counties sold 8.9% (t = -4.41) more stocks than managers in low-SCI counties. Importantly, social connections appear to be related to funds' selling activities during the pandemic; funds that are highly socially connected to hotspots sold 12.0% (t = -4.26) more of their holdings than their low-SCI counterparts. Furthermore, the effect of social connection on stock sales was elevated among what we call the "epicenter" stocks – stocks in industries that were most susceptible to the pandemic shock.

Owing to the highly exogenous nature of the pandemic outbreak, there is no reason to attribute the effects of hotspots and social connectedness to firm or fund features. First, COVID-19 was an unforecastable event prior to January 2020. Consequently, mutual funds are unlikely to make their headquarter locations based on their susceptibility to COVID-19. Furthermore, the experimental set up provides a natural control—low-SCI funds. Both the treatment (hotspot and high-SCI) and control groups were exposed to COVID-19; the main difference between the two is how socially connected these firms are to areas with high levels of COVID-19. Therefore, the differences in selling in these two groups should be related to social connectedness. This reasoning is borne out in the panel regression analysis; hotspot and high-SCI managers sold more shares even after controlling for geographical proximity to hotspots, case numbers, fund flows, economic exposure to COVID hotspots, other fund characteristics, a multitude of fixed effects, and a variety of firm characteristics.

The second part of our empirical tests aims to distinguish the smart connection hypothesis from the salience hypothesis. Even though social connection leads to more stock selling, the quarterly frequency of the fund holdings data does not pin down the exact timing of the selling. Therefore, we examine mutual fund performance during the second quarter of 2020 to distinguish whether social connections benefit or hurt investment performance. We do not begin the fund return examination from the first quarter because the COVID outbreak started in the middle of the first quarter in the USA and our hotspot and social connectedness measures are based on information at the end of the first quarter.

We find that the effect of social connections to COVID hotspots on fund manager behavior depended critically on manager skill. We measure manager skill using fund historical alpha (based on the CAPM model or the Carhart (1997) factors) or the Berk and van Binsbergen (2015) value added; managers are classified as high skilled if they are in the top 30% of the skill metric. We find that in 2020Q2 when the market rebounded from the first quarter low, the low-skill managers located in or socially connected to COVID hotspots had lower returns relative to the unconnected managers. A possible explanation for this finding is that social connections magnified the salience bias among unskillful managers and they thus timed their stock trading sub-optimally (e.g., sold out rebounding stocks too soon or got back to the market too late in the second quarter). On the other hand, high-skill managers experienced no net negative impact on their fund returns for being in a hotspot or being socially connected to a hotspot. Therefore, the overall evidence is in support of the salience hypotheses in that low skilled fund managers who were socially connected to the hotspots underperformed relative to their unconnected peers.

Our research contributes to the literature by exploring how geographical location and social connections influence institutional investor trading. On one hand, early research shows that fund managers prefer to make investments in local firms and firms that earn superior risk-adjusted returns (Coval and Moskwitz 1999, 2001) and later research links this to informal networks allowing the transfer of superior information (Cohen et al. 2008; Hong and Xu 2019; Bernile et al. 2015). In contrast, other recent research shows that these informal networks can increase the salience of extreme outcomes among retail investors, which encourages them to make high variance and skewness bets (Han et al. 2019; Bali et al. 2019). Our research extends both streams of the literature by focusing on the impact of social connections on fund manager behavior.

Two other recent papers also use the Facebook social connectedness index (SCI) data to study investor behavior. Bali et al. (2019) measure social connectedness of a stock's investor base and study the effect of social connectedness on lottery stocks. They find socially connected retail investors drive up lottery stock prices, leading to low returns of these stocks. Further, Kuchler et al. (2020) show that fund managers are more likely to invest in stocks that are more socially connected to them. Methodologically, our study differs from these two papers in how we use Facebook SCI data to study the impact of social connectedness on investment decisions. Bali et al. (2019) measures the social connection between a firm's headquarter county and all other counties. Kuchler et al. (2020) measures the social connection between the location of the fund manager and the location of the firm. In comparison, we measure social connectedness of fund managers to COVID-19 hotspots in 2020Q1 and examine how this social connection impacts managers' portfolio holdings. Since our SCI measure is premised on a well-defined theme (COVID), it clearly reflects the economic channel of the SCI effect.

In contrast to Kuchler et al. (2020) who find SCI-driven holdings do not differ from other holdings in investment performance, and to Cohen et al. (2008) who find network connections between fund managers and corporate board members help money managers gain an investment edge, we find that the effect of social connections on institutional investor behavior depends on investor skill: social connections to COVID hotspots exacerbated fears and suboptimal trading behaviors that hurt the performance of unskillful managers, while skillful managers appeared to have benefited from the informational advantage through connections with the hotspots during the pandemic. If retail investors are viewed as possessing generally low skills relative to institutional managers, our finding is conceptually consistent with Bali et al. (2019) who find that socially connected (retail) investors are more prone to salience bias and overbid lottery stocks. To our knowledge, our paper is the first to suggest that social connection helps intensify behavioral heuristics such as salience bias even for some institutional investors.

Our paper is also related to Alok et al. (2020), who examine the impact of geographical distance on the trading decision of professional money managers and find that fund managers located close to disaster cities irrationally underweight disaster zone stocks. However, our paper differs from Alok et al. (2020) in two ways. First, while Alok et al. conduct an event study of various natural disasters across time, we focus on COVID-19, an episode which has an immense impact on the economy in general, and an extraordinary influence on the role of SCI in investor behavior in particular. Second and more importantly, in contrast to Alok et al. who focus on the salience bias related to geographic location (proximity to disaster zone), we focus on social connectedness of investors to COVID hotspots. Therefore, we uncover a clear pathway through which the location of fund managers affects their trading decisions: it is the social connection to the source of "fear" (COVID), rather than the geographical distance to hotspots per se, that matters to investor perception of risk.[7]

2.2 Literature and Hypothesis Development

Our first two hypotheses link the location of institutional investors to stock selling during the pandemic outbreak period (2020Q1), without distinguishing the salience versus the smart connection hypotheses. The first hypothesis predicts how portfolio managers react during local

[7] We find that in our sample, the SCI effect persists even if we control for geographic distance between the fund manager and the hotspots, suggesting that geographical distance between two locations is a (negative) coarse measure of social connectedness between the people in the two locations.

exposure to COVID-19. Under the model of Bordalo et al. (2012), the salience of a payoff has a large impact on the decision makers' choice. Given that COVID-19 is more salient in cities with significantly higher cases and deaths, the model would predict that portfolio managers in these cities become much more risk averse than managers in non-hotspot cities. Other papers find evidence that salience matters in decision-making (Eraker and Ready 2015; Guiso et al. 2018; Dessaint and Matray 2017).

In terms of rational explanations, previous research finds that funds are locally biased and that these funds have superior information available to them (Coval and Moskowitz 1999, 2001). Given that COVID-19 has a deleterious effect on the real economy, being local to COVID-19 hotspots will arm managers with valuable information about firms affected by the pandemic and will trade accordingly during the pandemic outbreak.

H1: Portfolio managers in <u>COVID-19 hotspot counties</u> reduce stock holdings during the pandemic outbreak quarter (2020Q1).

The second hypothesis relates social connection to stock selling during the pandemic outbreak. Though the literature around social connectedness is still emerging, there are strong indications that social connections lead to similar trading behavior between locations. Several theoretical models predict that information sharing between traders have a large impact on trading which can cause momentum and other behavioral patterns (Duffie et al. 2009; Andrei and Cujean 2017). Recent empirical studies show that social connectedness matter for trading patterns such as those related to lotter stocks (Bali et al. 2019; Bali et al. 2011; Kuchler et al. 2020; Hirshleifer et al. 2020). Consequently, we expect both valuable information and irrational fear about COVID-19 in the local hotspot to be transferred to other localities via social connections, leading to strong stock selling from institutional investors in these localities.

H2: Portfolio managers in counties socially connected to COVID-19 hotspots reduce stock holdings during the pandemic outbreak quarter (2020Q1).

Our third hypothesis distinguishes the salience and rational theories about the effects of social connectedness. Under the salience hypothesis, social connections may result in behaviorally poor choices. Portfolio managers may focus on the potential extreme negative outcomes when presented with salient prospects of illness, death, or economic ruin from their social connections (Bordalo et al. 2012). This is further supported by other research that shows that portfolio managers overreact to natural and aviation disasters (Kaplanski and Levy 2010; Bernile et al. 2018; Alok et al. 2020). Finally, social connections may grant familiarity with another city, but not actually facilitate useful information transfer between two locales (Pool et al. 2012).

On the other hand, there is some evidence that institutional owners will use the information obtained from social connections to improve portfolio returns. For example, Coval and Moskowitz (2001) find that local portfolio managers outperform on local investments. Fund managers that are connected to a firm's operations in other states overweight stocks and demonstrated positive abnormal returns (Bernile et al. 2015).

Under the salience hypothesis, institutional investors located in or socially connected to the COVID hotspots oversell stocks out of fear during the pandemic outbreak, and trading under the influence of fear tends to be suboptimal. Under the smart connection hypothesis, institutional investors who are located in or socially connected to the hotspots make more informed trading decisions and tend to outperform their peers shortly after the pandemic outbreak.

These hypotheses should be mediated by manager skill. Previous research finds that less experienced managers suffer from an increase in home bias, without any corresponding increase in return (Pool, Stoffman, and Yonker 2012). Furthermore, lower skilled fund managers are associated with a higher degree of herding and lower returns (Jiang and Verardo 2018). Consequently, we expect that the salience or smart connection hypotheses will depend on fund manager skill. Less skilled managers will suffer more from salience bias, and thus make poor trades, while in contrast, more skilled managers will suffer less from this bias. We therefore have two opposite predictions:

H3a: According to the salience hypothesis, fund managers located in or socially connected to the hotspots, especially those with low skills, underperform their peers in the quarter following the pandemic outbreak (2020Q2).

H3b: According to the smart connection hypothesis, fund managers located in or socially connected to the hotspots, especially those with high skills, outperform (or underperform less) relative to their peers in the quarter following the pandemic outbreak (2020Q2).

2.3 Data

Our data come from several sources. For COVID-19 cases and deaths, we use the New York Times (NYT) county-level data.[8] This data is collected by NYT's journalists by combing through news conferences, data releases, and speaking with public officials. We assumed that any US county with no data had 0 cases and 0 deaths and that cases and deaths in New York City, which was reported as a single entity, were evenly split across its five counties. We then used the city-to-county finder in SAS to link county information to cities; this is required because mutual fund headquarter data is by city while COVID data is by county.

[8] https://raw.githubusercontent.com/nytimes/covid-19-data/master/us-counties.csv (last accessed July 31, 2020)

Our second dataset is the Facebook Social Connectedness Index (*SCI*). It provides a measure of the amount of Facebook friendship connections between one US county and another (see Appendix B for a full description of Facebook's *SCI* methodology). The score is best framed as the relative probability of a Facebook friendship given the two users counties. As above, we used the city-to-county finder to link the county data to cities. The dataset we used in this paper is from March 2020 and thus represents an accurate snapshot of connections for the pandemic and 2020Q1 institutional ownership reporting period.

The third source for mutual fund and institutional ownership data is the Center for Research in Security Prices Mutual Fund (CRSP MF) data. We screen funds by a process similar to Hong et al. (2005). We remove funds a) outside of the US, b) with less than 10% stock holdings, c) index funds,[9] and d) funds that only report semi-annually. Furthermore, we only include funds that are all-equity—mixed bond and stock funds are excluded—to ensure that the results are not due to shifting between bonds and stocks. The database also provides addresses for the mutual fund headquarters and we use this information to determine which county and city the fund resided. Finally, we draw stock price data from CRSP.

These datasets are then used to derive key variables used in our analysis. The first variable is to define counties that were badly affected by COVID-19 as of March 30, 2020. This date was chosen to prevent look ahead bias. COVID-19 data was typically announced at the end of the day; consequently, March 31st data would not have been available to investors until after the close of the trading day. We flag counties with at least 2,000 cumulative cases as *Hotspot* = 1 and other counties as *Hotspot* = 0. Table 2.1 Panel A provides additional details about the hotspot counties.

[9] Drop funds with "index" in the name or funds that are not "active" in the "investment orientation" field.

Table 2. 1: Summary Information about COVID-19 Hotspot Counties and Epicenter Industries

This table shows the list of the Hotspot counties and Epicenter industries. Panel A provides the detailed information about the identified hotspot counties including the FIPS code, county name, associated city name, state name, cumulative number of cases and cumulative number of deaths as of March 30, 2020. The county level data is collected by New York Times journalists by combing through news conferences, data releases, and speaking with public officials. We assume that any US county with no data had 0 cases and 0 deaths and that cases and deaths in New York City, which is reported as a single entity, were evenly split across its five counties. Panel B shows the list of epicenter industries that had the largest equal-weighted decline in stock returns during 2020Q1. These industries use the Fama-French 48 industry definitions.

Panel A: Hotspot Counties

No.	FIPS	County	City	State	Cases	Deaths
1	36119	Westchester County	Purchase	NY	9326	19
2	36061	New York County	New York	NY	7617	183
3	36059	Nassau County	Syosset	NY	7344	48
4	36103	Suffolk County	Syosset	NY	5791	44
5	17031	Cook County	Chicago	IL	3727	44
6	6037	Los Angeles County	Los Angeles	CA	2474	44
7	53033	King County	Seattle	WA	2332	152

Panel B: Epicenter Industries

No.	Industry Name	Industry Description	FF48 Industry	EW Return (%)
1	Oil	Petroleum and Natural Gas	30	-67.91
2	Coal	Coal	29	-38.09
3	Clths	Apparel	10	-30.99
4	Autos	Automobiles and Trucks	23	-29.47
5	Aero	Aircraft	24	-27.17
6	Meals	Restaurants, Hotels, Motels	43	-25.64
7	Steel	Steel Works	19	-25.37
8	Txtls	Textiles	16	-24.93
9	Mines	Non-Metallic and Industrial Metal Mining	28	-24.8
10	Fun	Entertainment	7	-24.44
11	Ships	Shipbuilding, Railroad Equipment	25	-23.96
12	FabPr	Fabricated Products	20	-22.06
13	Rtail	Retail	42	-21.44
14	Chems	Chemicals	14	-21.32
15	Mach	Machinery	21	-18.57
16	Hshld	Consumer Goods	9	-17.08
17	Banks	Banking	44	-16.97
18	Trans	Transportation	40	-16.88
19	Paper	Business Supplies	38	-16.09
20	PerSv	Personal Services	33	-15.79
21	RlEst	Real Estate	46	-13.48
22	ElcEq	Electrical Equipment	22	-13.29
23	Books	Printing and Publishing	8	-13.15
24	Cnstr	Construction	17	-13.13

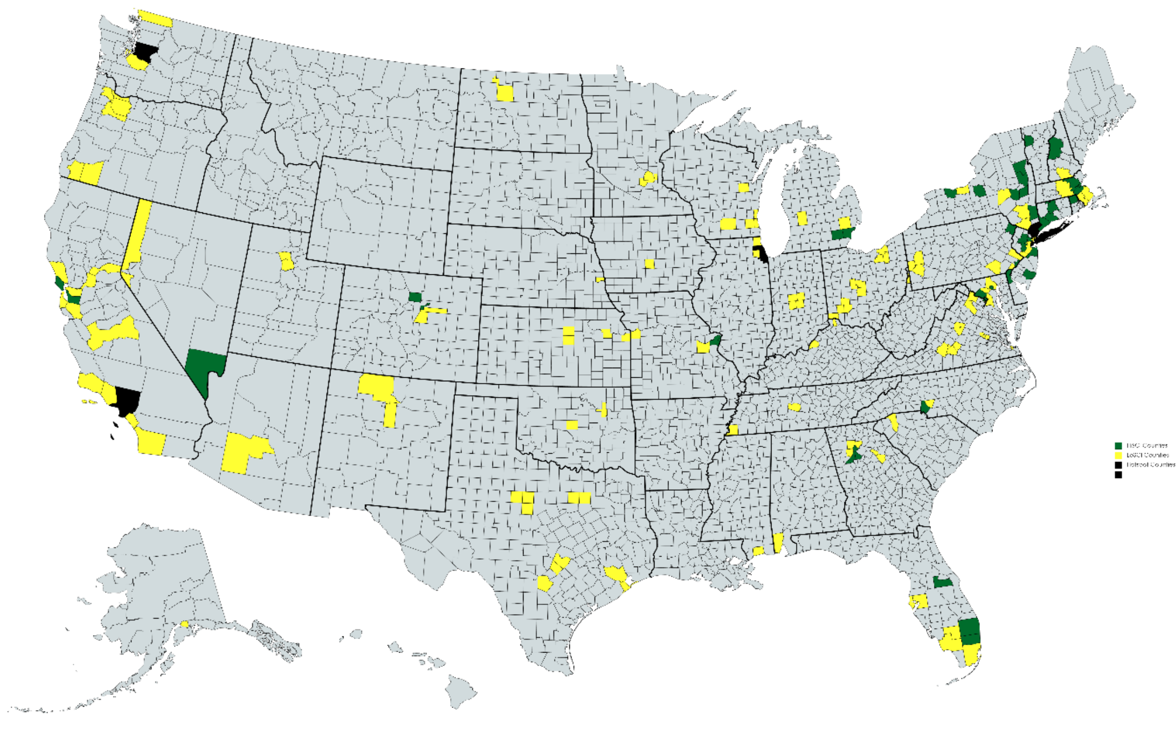
The second key variable is *HiSCI*, counties that are socially connected to counties facing high levels of COVID-19. First, we calculate each county's average *SCI* score to all of the *Hotspot* counties. Second, we flag a county as *HiSCI* = 1 if the county is in the top quartile of average *SCI* scores from the previous step and is *not* a hotspot county. It is otherwise set to 0 and vice versa for *LoSCI*. See Appendix 6.B for more details.

Figure 2.1 displays the geographic distribution of *Hotspot*, *HiSCI*, and *LoSCI* counties. These counties are broadly distributed across the US, suggesting that the *HiSCI* counties contains distinct information content from geographical proximity to hotspot counties. For example, several areas near San Francisco and Las Vegas are *HiSCI* countries largely because of their high social connections to the hotspot city of Los Angeles, even though they are geographically farther away from Los Angeles than some *LoSCI* countries. However, we control for geographic proximity when testing for social connection effects in the regressions.

We use three measures of fund manager skill. The Carhart-4 alpha and CAPM alpha are more traditional metrics of average excess returns after controlling for the risks fund managers took (e.g., Barber et al. 2016). The Carhart-4 alpha is calculated for each fund manager over a 5-year rolling window, where we regress each fund manager's returns over the window against the Carhart (1997) 4 risk factors to find alpha (the excess return). We then rank all fund managers based on their alpha from highest to lowest and identify the top tercile (30%) of fund managers as having high skill; we use terciles to ensure we have adequate sample size of high-skilled managers and to avoid results being driven by outliers. We use an indicator variable (*Perf_Car4*) to define manager skill. Our results are robust to using the CAPM model (*Perf_CAPM*) instead of the Carhart 4 factors to estimate alpha.

Figure 2. 1: Geographical Distribution of the COVID-19 Hotspot, High and Low Social Connectedness Counties

This figure illustrates the geographic distribution of different categories of counties on the map of the United States. The counties filled with black color are COVID-19 *Hotspot* counties, which are defined as the counties that have more than 2000 cases by March 30, 2020. The counties filled with green and yellow are *HiSCI* and *LoSCI* counties, which are defined as non-*Hotspot* counties whose aggregate *SCI* to *Hotspot* counties are above (below) top quartile. The rest of the counties, with grey color, are those with no qualified fund headquarters.



The Berk and van Binsbergen (2015) metric is based on mutual fund value added, which they show is better at predicting future fund performance than historical alpha. To calculate this measure, we first take the return of each fund and add in management fees to construct month-*t* gross return. Next we subtract the Vanguard index fund return for each month *t*.[10] We then multiply this excess return by the assets under management from month *t*-1 to determine the value added that month. Third, we *t*-test each fund's value added over the pre-sample period to determine how much value firms add over the pre-sample period. The pre-sample period is the period between 2010 and the year preceding the date being tested; i.e., end date of 2018 for any 2019 mutual fund regressions. We exclude any fund that has less than 24 months of data in the pre-sample period.[11] Finally, we rank these firms from highest to lowest *t*-value and sort them into terciles based on skill. Funds in the top tercile are defined as high skill.

2.4 Empirical Results

2.4.1 Summary Statistics and Univariate Tests

Table 2.2 shows the summary statistics for fund holdings, *SCI*, and COVID data over the sample period. Panel A and B provides basic summary statistics for the funds from 2016Q1 to 2020Q1 and 2020Q1, respectively. It is clear from the 2020Q1 holding data, overall there was a large drop in the number of shares across funds. However, Panel C demonstrates that this drop in the number of shares over 2020Q1 was highly concentrated in *Hotspot* and *HiSCI* counties.

[10] As not all Vanguard index funds are available for the entire data period, we do a linear projection of the ith active mutual fund value onto the set of Vanguard index funds.
[11] The funds in our sample area all actively managed all-equity funds, so this already matches Berk and van Binsbergen (2015).

Table 2. 2: Summary Statistics

This table reports the summary statistics for the key variables. The variables are defined in Appendix A. Panel A includes the full sample between 2016Q1and 2020Q1. Panel B only includes the COVID-19 sample period of 2020Q1. Panel C includes the comparison of the summary statistics between the COVID-19 hotspot counties and the low-SCI counties and between high and low-SCI counties. Diff (*t*-stat) provides the two-sample *t*-test for the difference in mean. The epicenter industries (denoted by "*Epic*") are defined as the 10 Fama-French 48 industries with the lowest equal-weighted industry returns during 2020Q1.

Panel A: Full Sample Period (2016Q1 to 2020Q1)

	N	Mean	Std.Dev	P25	Median	P75
Fund AUM ($Million)	52,533	2,017.42	7,736.27	65.10	308.20	1,272.40
No. Stocks Held by Fund	52,533	105.36	236.65	21.00	48.00	95.00
Avg $ Invested per Stock ($Million)	52,533	68.52	497.37	1.43	6.74	29.73
No. Funds in County	154	21.97	65.06	1.00	4.00	21.00
Fund Age (Years)	3,383	16.57	13.04	6.50	14.25	23.33
Flow (%)	52,533	0.27	15.79	-4.58	-1.70	1.71
Fund_Expense	52,533	0.01	0.00	0.01	0.01	0.01
Fund_Mgt_Fees (%)	52,533	0.29	3.15	0.48	0.71	0.88
Fund_Turnover	52,533	0.79	1.99	0.27	0.48	0.82
Share Change (%)	5,023,657	-0.67	66.64	-12.11	0.00	3.83

Panel B: COVID-19 Sample Period (2020Q1)

	N	Mean	Std.Dev	P25	Median	P75
Fund AUM ($Million)	3,336	1,711.79	6,978.45	50.60	223.30	969.50
No. Stocks Held by Fund	3,336	111.87	258.55	22.00	48.00	95.00
Avg $ Invested per Stock	3,336	51.06	236.09	1.06	4.79	22.11
No. Funds in County	154	21.66	64.20	1.00	4.00	21.00
Fund Age (Years)	3,336	16.66	13.07	6.58	14.33	23.33
Flow (%)	3,336	-2.17	8.79	-5.24	-2.48	0.41
Fund_Expense	3,336	0.01	0.00	0.01	0.01	0.01
Fund_Mgt_Fees (%)	3,336	0.35	3.07	0.44	0.67	0.85
Fund_Turnover	3,336	0.78	2.00	0.26	0.46	0.80
No. Cases in County	154	533	1,268	61	185	396
No. Deaths in County	154	8.21	21.69	1.00	2.00	6.00
Avg SCI to Hotspot County	154	2,426	1,073	1,602	2,230	3,186
Share Change (%)	347,729	-4.25	64.96	-21.76	0.00	9.72
Share Change (%) – *Epic*	37,790	-10.18	70.89	-48.18	0.00	7.71

Panel C: Comparisons of *Hotspot* and *HiSCI* Counties Relative to *LoSCI* Counties, 2020Q1

	***Hotspot* Counties**		***HiSCI* Counties**		***LoSCI* Counties**		***Hotspot - LoSCI***	***HiSCI- LoSCI***
	N	**Mean**	**N**	**Mean**	**N**	**Mean**	**(*t*-stat)**	**(*t*-stat)**
Fund AUM ($Million)	953	2016.98	950	1944.70	1433	1357.29	(1.74)	(2.67)
No. Stocks Held by Fund	953	97.55	950	93.21	1433	133.76	(-3.34)	(-3.94)
Avg $ Invested per Stock ($Million)	953	50.89	950	58.44	1433	46.29	(-0.48)	(-1.11)
No. Funds in County	7	136.14	38	25	109	13.15	(1.55)	(0.93)
Fund Age (Years)	953	16.28	950	17.44	1433	16.4	(-0.22)	(1.88)
Flow (%)	953	-3.02	950	-1.95	1433	-1.75	(-3.15)	(-0.55)
Fund_Expense	953	0.01	950	0.01	1433	0.01	(-2.81)	(-2.81)
Fund_Mgt_Fees (%)	953	0.15	950	0.47	1433	0.4	(-1.5)	(0.93)
Fund_Turnover	953	0.84	950	0.72	1433	0.78	(0.78)	(-0.62)
No. Cases in County	7	5515.91	38	487.71	109	228.29	(5.11)	(3.23)
No. Deaths in County	7	76.26	38	8.76	109	3.64	(3.02)	(2.05)
Avg SCI to Hotspot County	7	3590.82	38	3798.7	109	1872.26	(4.14)	(16.24)
Share Change (%)	82022	-8.45	85513	-12.19	180194	1.36	(-33.64)	(-51.88)
Share Change (%) – *Epic*	9411	-14.09	8971	-21.89	19408	-3	(11.02)	(21.48)

Table 2. 3: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q1

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between 2016Q1 and 2020Q1. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change (%)*	(2) *Share Change (%)*	(3) *Share Change (%)*	(4) *Share Change (%)*	(5) *Share Change (%)*	(6) *Share Change (%)*
Hotspot	-1.040	-1.334	1.184	0.972		
	(-0.92)	(-1.12)	(1.14)	(0.91)		
Hotspot×COVID	-8.778***	-10.104***	-9.214***	-8.818***	-9.152***	-8.859***
	(-3.44)	(-3.95)	(-4.13)	(-4.00)	(-4.49)	(-4.41)
HiSCI	-0.029	0.284	-0.907	-1.245		
	(-0.03)	(0.28)	(-1.12)	(-1.51)		
HiSCI×COVID	-13.522***	-12.079***	-10.368***	-10.010***	-12.092***	-11.982***
	(-4.23)	(-4.01)	(-3.78)	(-3.66)	(-4.25)	(-4.26)
GEO		-1.596	0.293	0.034		
		(-1.59)	(0.35)	(0.04)		
GEO×COVID		-6.943*	-5.503	-5.177	-5.780	-5.468
		(-1.72)	(-1.45)	(-1.37)	(-1.50)	(-1.42)
COVID	1.492	2.819	3.277*		2.491	
	(0.85)	(1.60)	(1.89)		(1.54)	

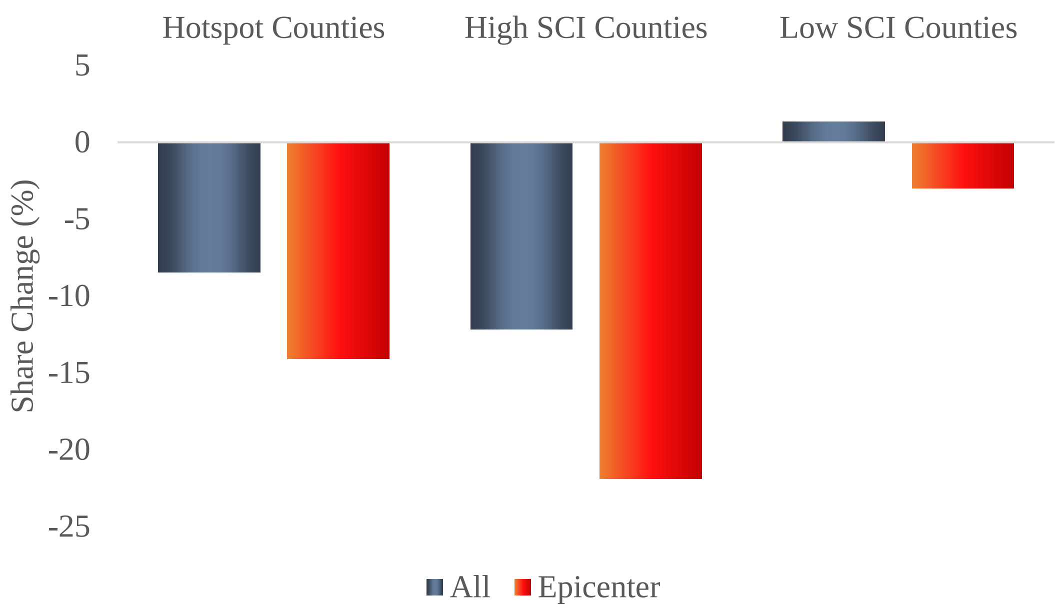
Figure 2.2 provides a visualization of how fund managers in *Hotspots*, *HiSCI*, and *LoSCI* counties acted differently. In the blue columns, managers in *Hotspot* had a mean percentage decline in stock holdings of -8.45% and managers socially connected to *Hotspot* regions had a mean decline of -12.19%. In comparison, socially unconnected managers experienced a 1.36% increase in stock holdings during Q1 2020. Furthermore, the intensified selling concentrates in the epicenter stocks, or stocks in industries with the lowest equal-weighted returns in 2020Q1 (see Table 2.1 Panel B for details of the epicenter industries). The red columns show that funds in the hotspots, high-SCI, and low-SCI counties sold 14.09%, 21.89%, and 3.00% of their epicenter stock holdings, respectively. Panel C of Table 2.2 indicates that the levels of stock sales of the hotspot and high-SCI funds are highly significantly different from that of the low-SCI counterparts.

2.4.2 Impact of COVID-19 Hotspots and Social Connectedness on Stock Holdings

While the univariate results in Table 2.2 provide an excellent first test of the differences for funds in *Hotspot* and *HiSCI* counties there needs to be additional controls for the observable and unobservable differences in the funds. For example, Table 2.2 Panel C shows that funds in *HiSCI* and *LoSCI* counties have differing levels of COVID-19 with 488 versus 228 cases (t = 3.23), respectively. Further, *HiSCI* funds hold fewer unique stocks and invest more per stock. Consequently, we will use a combination of firm, fund, quarter, and fund × industry fixed effects to help control for these observable and unobservable variables. We also include controls for fund characteristics, and COVID exposure such as geography, number of cases, or economic linkages to COVID hotspots to control for any differential infection rates of COVID-19 during the 2020Q1 period.

Figure 2. 2: Percentage Change in Shares over 2020Q1, for Funds in COVID-19 Hotspots, High and Low Social Connectedness Counties

This figure displays the firm-fund level average of the percentage change in shares (% Change in Shares). *Hotspot* counties are defined as counties that have more than 2,000 COVID-19 cases by March 30, 2020. High *SCI* (Low *SCI*) counties are defined as non-hotspot counties whose aggregate *SCI* to hotspot counties are above (below) the top tercile.



$$\mathit{\Delta}H_{i,j,q} = \alpha + \beta_1\, Hotspot + \beta_2\, Hotspot \times COVID + \gamma_1\, HiSCI + \gamma_2\, HiSCI \times COVID$$

$$+ \theta_1\, GEO + \theta_2\, GEO \times COVID + \vartheta_i + \pi_j + \tau_q + \delta_{j,Ind} + \varepsilon_{i,j,q}, \tag{2.1}$$

where $\mathit{\Delta}H_{i,j,q}$ is the core dependent variable, percentage change in shares (*Share Change %*) for firm i, fund j, and quarter t. *Hotspot*, *HiSCI*, and *GEO* are indicator variables equal to one when the mutual fund's headquarter located in hotspot counties, non-hotspot counties that are highly socially connected to hotspots, and non-hotspot counties that are geographically proximate to hotspots, respectively. *COVID* is an indicator variable which equals one when it is 2020Q1. ϑ_i, π_j, τ_q, and $\delta_{j,Ind}$ are controls for the firm, fund, quarter, and fund × industry fixed effects. We omit other controls in the equation for brevity.

We include a wide variety of fund characteristics as controls to ensure that the results are not due to other causes such as investor flows, fund size, fund fees, or fund trading behavior. *Flow* controls for investor flows over the quarter to ensure that the results are not driven by a large influx (or outflow) of investor funds. *LnAUM* and *LnAge* controls for fund size and age while *Fund_Mgt_Fees* and *Fund_Expense* control for any fund fees. Finally, *Fund_Turnover* controls for funds likelihood to buy or sell shares while $VRet_{q-1}$ controls for the previous quarter's fund performance.

Furthermore, the regressions control for funds' economic exposure to COVID to ensure that the results account for funds' COVID exposure in their portfolios. To control for proximity to a COVID hotspot, we include *GEO*, an indicator variable set to 1 if it is within 100 miles of a hotspot. This ensures the results are not driven by funds close to large outbreaks of COVID-19. Moreover, the regression controls for pre-existing economic linkages to COVID hotspots proxied by

HotspotCorp. This variable is defined as the weight of a fund's portfolio that is invested in firms headquartered in a COVID hotspot in *q*-1.

The first step in our analysis is to confirm that high levels of COVID-19 in a county leads to a reduction in stock holdings by fund managers. According to hypothesis H1, close exposure to COVID-19 has both a real negative economic effect on the local economy as well as greatly increasing the salience of the destructive power of the disease. Consequently, fund managers in counties with high levels of COVID-19 should reduce their holdings of risky assets, such as stocks.

To test the impact of being in an area of high levels of COVID-19, we focus on the indicator variable *Hotspot* and its interaction with *COVID*. In columns 1-4, the *Hotspot* variable by itself does not show any significance outside of the COVID period (fund fixed effects absorb the *Hotspot* effect in columns 5-6). In all columns, the coefficient on *Hotspot* × *COVID* exhibits a significant hotspot effect on stock selling during the COVID outbreak. This supports our first hypothesis that direct exposure to COVID-19 will reduce funds' share ownership.

The second step in our analysis is to test hypothesis H2: whether mutual fund managers with a high social connectedness to a COVID-19 hotspot influences their stock holdings. Table 2.3 shows that *HiSCI* is significantly negatively related to the percentage change in the number of shares held (columns 1). Further, the negative significant relationship remains even after controlling for geographic distance to hotspots (column 2), fund flows, and the percentage of funds' portfolio in headquartered in COVID hotspots, and other controls (columns 3 and 4). Finally, the social connection to hotspots continues to be significant even if fund and fund × industry fixed effects are used (columns 5 and 6).

Table 2.3: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q1 (Continued)

	(1) *Share Change (%)*	(2) *Share Change (%)*	(3) *Share Change (%)*	(4) *Share Change (%)*	(5) *Share Change (%)*	(6) *Share Change (%)*
Flow			0.940***	0.943***	0.953***	0.961***
			(20.15)	(20.00)	(18.63)	(18.51)
HotspotCorp			0.028	0.009	-0.468*	-0.434*
			(0.49)	(0.14)	(-1.70)	(-1.69)
LnAge			-1.247**	-1.248**	-11.751***	-2.016
			(-2.34)	(-2.30)	(-8.74)	(-1.13)
LnAUM			1.226***	1.238***	-1.487	-1.398
			(4.74)	(4.79)	(-0.95)	(-0.99)
Fund_Mgt_Fees			-0.777***	-0.710***	-0.221	-0.367
			(-3.22)	(-2.80)	(-0.97)	(-1.47)
Fund_Expense			-2.815*	-3.277*	11.155**	2.231
			(-1.78)	(-1.94)	(2.32)	(0.47)
Fund_Turnover			-0.398	-0.297	0.002	-0.024
			(-1.26)	(-0.93)	(0.01)	(-0.09)
$VRet_{q-1}$			0.019	-0.066	0.016	-0.108
			(0.68)	(-0.95)	(0.57)	(-1.58)
Observations	5,023,657	5,023,657	4,876,486	4,876,295	4,836,605	4,836,418
Adjusted R^2	0.001	0.001	0.037	0.050	0.059	0.074
Firm FE	No	No	No	Yes	No	Yes
Fund FE	No	No	No	No	Yes	Yes
Quarter FE	No	No	No	Yes	No	Yes
Fund × Industry FE	No	No	No	No	Yes	Yes

Based on the coefficients on *Hotspot* × *COVID* and *HiSCI* × *COVID* in the full model in Table 2.3 using fund and fund × industry fixed effects (column 6), fund managers in a hotspot county reduced stock holdings by -8.86% relative to the control group (low-SCI managers). Similarly, managers with high social connection with the hotspots unloaded holdings by an additional -11.98% relative to their low-SCI counterparts. Therefore, the economic impact of social connectedness is at least comparable to that of the hotspot itself.

The effects of hotspot and social connectedness during the COVID-19 outbreak appear stronger than those documented in other studies using more general samples, but this is likely due to the size of the COVID-19 impact. For example, Alok et al. (2020) find that being close to a disaster zone reduced fund managers' disaster zone stock holdings by 1.5%; Kuchler et al. (2020) document that a 10% increase in social connectedness between firm and investor locations is associated with an increase in stock weight by 1.9% in the investor's portfolio. However, the larger effect is intuitive given the impact of the COVID-19 pandemic on the US economy. Furthermore, the COVID-19 pandemic is a highly salient event and social interactions have especially strong effects on investors when transmitting salient information (Han et al. 2019; Hirshleifer et al. 2020).

One possible alternative explanation of our results is that funds with stronger social connections with COVID hotspots are also more economically connected to the hotspots, causing more intensive stock selling of *HiSCI* managers. We address this concern using two controls for economic ties. First, we use a proxy, *Hotspotcorp*, defined as the weight of a fund's portfolio that is invested in firms headquartered in a COVID hotspot in *q*-1, to capture the economic link between a fund and the hotspots. Table 2.3 shows some evidence that this variable is negatively related to our measures of share change, consistent with the interpretation that economic links lead to stock selling during the viral outbreak.

Second, one way economic ties influence investor stock selling is fund redemptions: it is possible that investors who had strong ties with COVID hotspots suffered heavy losses from assets related to the hotspots, forcing them to redeem funds and leading to heavy stock selling of *HiSCI* funds. We use fund flows (*Flow*) to measure redemptions, and find it has a strong positive relation with various measures of share change, suggesting that redemptions, as reflected by the negative fund flows during 2020Q1 (Table 2.2, Panel B), indeed contributed to stock selling. However, the *HiSCI* effect persists after controlling for economic links.

Types of Stocks Sold

A further test of our results is to examine what types of stocks funds sell. Figure 2.2 provides preliminary evidence that the social connection effect on stock selling is greater among stocks most heavily hurt by the pandemic. We now provide multivariate evidence.

In Table 2.4, we subdivide the stocks into "epicenter" (*Epic*) stocks, or stocks most hurt by the pandemic in 2020Q1. We define *Epic* stocks to be those in the 10 (or 24) most underperforming industries (measured by equal-weighted returns during 2020Q1) of the Fama-French 48. As can be seen from the coefficient of *HiSCI* × *COVID* × *Epic* in columns 1-4, socially connected funds sold more *Epic* stocks than non-*Epic* stocks. This confirms that fund managers reacted to the crisis by selling the stocks that were most hard-hit in the early part of the pandemic. Interestingly, judging by the coefficient of *Hotspot* × *COVID* × *Epic,* the effect of *Hotspot* on stock sales are not intensified among epicenter stocks, lending further credence to the conclusion that the effect of social interaction on investor behavior is distinct from that of geographical location.

Table 2. 4: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors: Epicenter Stocks

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between 2016Q1 and 2020Q1. The regression includes an interaction for *Epic*, set to 1 if the stock is in the 10 (or 24) most underperforming industries (measured by equal-weighted returns during 2020Q1) of the Fama-French 48, and 0 otherwise. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change %*	(2) *Share Change %*	(3) *Share Change %*	(4) *Share Change %*
Hotspot	1.492		1.865**	
	(1.52)		(2.00)	
Hotspot×COVID	-8.909***	-8.965***	-8.714***	-8.601***
	(-4.04)	(-4.52)	(-4.17)	(-4.56)
Hotspot×COVID×Epic	-0.334	-0.695	-0.778	-1.438
	(-0.17)	(-0.34)	(-0.72)	(-1.32)
HiSCI	-0.752		-0.407	
	(-0.98)		(-0.55)	
HiSCI×COVID	-10.155***	-11.870***	-9.815***	-11.164***
	(-3.60)	(-4.09)	(-3.52)	(-3.91)
HiSCI×COVID×Epic	-4.988***	-6.047***	-2.585**	-3.959***
	(-2.94)	(-3.37)	(-2.31)	(-3.20)
GEO	0.389		0.526	
	(0.48)		(0.68)	
GEO×COVID	-4.911	-5.022	-4.750	-4.897
	(-1.23)	(-1.26)	(-1.18)	(-1.23)
GEO×COVID×Epic	-2.722	-3.353*	-1.267	-1.423
	(-1.61)	(-1.91)	(-1.16)	(-1.22)
Epic	-0.067		0.813***	
	(-0.20)		(2.83)	
COVID	3.402**		3.818**	
	(1.99)		(2.36)	
Observations	4,437,292	4,397,707	4,437,292	4,397,707
Adjusted R^2	0.038	0.076	0.038	0.076
Other *Epic* Interactions	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Quarter FE	No	Yes	No	Yes
Fund × Industry FE	No	Yes	No	Yes
Epic Industry	Lowest 10 EW	Lowest 10 EW	Lowest 24 EW	Lowest 24 EW

2.4.3 A Question of Skill: Rational or Behavioral Fund Manager Reaction

The third hypothesis is a dual sided one. On one hand, fund managers have historically overweighted investments in local firms, but earned superior returns due to better information—social connectedness may thus improve information available to fund managers. On the other hand, social interactions could enhance the salience of the economic and health problems of COVID-19 relative to fund managers with low social connections with the hotspots.

We modify the regression equation (2.1) from above by changing the dependent variable to the future aggregate return for fund *j* at quarter *q*+1 ($VRet_{j,q+1}$). We also change the fixed effects to fund style × quarter fixed effects; fund styles are from the CRSP MF database which uses the Lipper Fund Classification. The control variables are similar to equation (2.1) and adjust for fund characteristics that may drive fund performance (such as fund flows, fund size, or fees) and the fund exposure to COVID-19 (such as geographic proximity to hotspots, COVID cases in its county, or percentage of portfolio economically linked to the hotspots):

$$VRet_{j,q+1} = \alpha + \beta_1\, Hotspot + \beta_2\, Hotspot \times COVID + \gamma_1\, HiSCI + \gamma_2\, HiSCI \times COVID$$

$$+ \theta_1\, GEO + \theta_2\, GEO \times COVID + \pi_j \times \tau_q + \varepsilon_{j,q} \qquad (2.2)$$

where $\pi_j \times \tau_q$ indicates fund style × quarter fixed effects. If social connections are associated with useful information, we should expect funds highly socially connected to hotspots (*HiSCI*) to outperform during the post COVID outbreak period (*HiSCI × COVID)*. If the social connections transmit salient fears instead, we should expect the opposite— *HiSCI × COVID* should have a negative coefficient.

To examine this, in Table 2.5 we look at forward fund returns (*VRet*$_{q+1}$), with 2020Q2 being our focus period. Interestingly, in non-COVID periods, socially connected funds seem to outperform their unconnected funds slightly for future returns. This suggests that these connections provide useful information during normal periods. For example, *HiSCI* fund managers outperformed an average of 14–22 basis points (bp) in non-COVID periods (columns 1-2). Possibly, fund managers' social connections to hotspots provide useful information between fund managers in *HiSCI* areas in normal (non-COVID) circumstances. This provides some support for the smart connection hypothesis.

On the other hand, fund social connections seem to hurt in the future during periods of extreme stress; *HiSCI* × *COVID* has a negative coefficient for columns 1 and 2. Based on the column 2 results, a socially connected fund underperformed by 30.8 bp (t = -4.60) during the pandemic relative to a non-connected fund, which swamps the benefits of *HiSCI* during non-COVID periods (14.0 bp). The salience bias transferred to the fund managers causes them to underperform in the quarters subsequent to the viral outbreak, possibly because the managers let their fears delay them from putting money back to the market.

Conditioning on Manager Skill

To test Hypothesis 3 more fully, we condition our tests on manager skill. We make use of three different versions manage skill metrics, which we interact with key variables in equation (2.2). In the first two versions, *Perf_CAPM/Perf_Car4* is an indicator variable is set to 1 if a fund was in the top tercile of alpha in the past 5 years after adjusting for the CAPM or Carhart (1997) 4 factors. In the third version, *Perf_BB* is an indicator variable set to 1 if a fund was in the top tercile of performance according to Berk and van Binsbergen (2015). This approach will let us see how the effect of social connection on fund performance depends on manager skill.

Table 2. 5: Fund Future Quarterly Return Regressions

This table provides the results of the panel regression of future fund returns (quarterly returns over the next quarter) between 2016Q2 and 2020Q2. The dependent variable $VRet_{q+1}$ is the value-weighted return of the fund less fees in q+1. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Control variables are identical to those in Table 2. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and quarter, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) $VRet_{q+1}$	(2) $VRet_{q+1}$
Hotspot	0.111	0.081
	(1.70)	(1.39)
Hotspot×COVID	-0.155**	-0.136*
	(-2.56)	(-2.02)
HiSCI	0.217**	0.140*
	(2.63)	(1.91)
HiSCI×COVID	-0.352***	-0.308***
	(-10.65)	(-4.60)
GEO	0.003	0.072
	(0.03)	(0.57)
GEO×COVID	-0.475***	-0.399***
	(-8.66)	(-3.36)
Flow		0.003
		(1.24)
HotspotCorp		0.021***
		(4.00)
LnAge		0.058
		(1.06)
LnAUM		0.026
		(0.84)
Fund_Mgt_Fees		-0.032
		(-1.31)
Fund_Expense		-8.411
		(-0.47)
Fund_Turnover		0.009
		(0.18)
$VRet_{q-1}$		1.928
		(0.23)
Observations	53,458	49,723
Adjusted R^2	0.794	0.801
Fund Style × Quarter Fixed Effect	Yes	Yes

The results in Table 2.6 demonstrate a strong dichotomy of the effects of social connections on fund performance depending on manager skill. Judging from *HiSCI* × *COVID* in columns 2, 4, and 6, low-skilled managers socially connected to hotspots experienced large losses of -66.5 bp, -49.8 bp, and -54.2 bp (*t*-value -6.84, -5.76, and -6.11, respectively) based on CAPM alpha, Carhart-4 alpha, or Berk and van Binsbergen (BB2015) metrics for skill. This shows that regardless of whether we measure fund manager skill based on historical alpha from the CAPM or Carhart-4 models or from the BB2015 metric, unskilled managers suffered from salience bias from being in or socially connected to the COVID disaster.

On the other hand, for high-skill managers, being socially connected to COVID hotspots helps performance in the next quarter. Based on the coefficient of *HiSCI* × *COVID* × *{Perf}* in columns 2, 4, and 6, relative to the unskilled managers, we see that the skilled managers outperform by 83.3 bp, 71.9 bp, and 69.0 bp (*t*-values 7.87, 5.63, and 5.25) in the quarter after the initial COVID outbreak, respectively.[12] Therefore, higher skilled managers were better able to survive the COVID turmoil than lower skilled managers. The results are similar for high skilled fund managers in the hotspots with the coefficients for *Hotspot* × *COVID* × *{Perf}* all being positive and significant for columns 1 and 3-6.

Moreover, there is evidence that the high skilled managers outperformed even relative to unconnected managers during the COVID period. The sum of *HiSCI* × *COVID* + *HiSCI* × *COVID* × *{Perf}*, which measures the net effect of *HiSCI* of high skilled managers on 2020Q2 performance relative to unconnected managers, is 16.8 bp, 22.2 bp, and 14.8 bp (*t*-values 1.68, 2.25, and 1.31) in columns 2, 4, and 6, respectively. Similarly, the sum of *Hotspot* × *COVID* + *Hotspot* × *COVID*

[12] Note these percentage gains and losses seem large in isolation; however, there was a tremendous amount of market movement in this period. The S&P 500 index fell 20.67% in 2020Q1 and rose 18.13% in 2020Q2.

Table 2. 6: Fund Future Quarterly Return Regressions, Conditioning on Fund Manager Skill

This table reports results of the panel regressions of future fund returns ($VRet_{q+1}$, quarterly returns over the next quarter) between 2016Q1 and 2020Q1, conditioning on fund manager's CAPM alpha in the past 5 years (col 1-2), Carhart 4 alpha in the past 5 years (col 3-4), and fund manager skill calculated as per Berk and van Binsbergen (2015) (col 5-6). The dependent variable is the value-weighted return of the fund less fees in quarter q+1 ($VRet_{q+1}$). *Perf_CAPM/Perf_Car4/Perf_BB* is an indicator variable set to 1 if the fund is in the top-tercile of CAPM alpha, Carhar-4 alpha, or Berk and van Binsbergen (2015) manager skill and 0 otherwise. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Control variables are identical to those in Table 2. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and quarter, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

Skill Measure *{Perf}*	*Perf_CAPM*		*Perf_Car4*		*Perf_BB*	
	(1)	(2)	(3)	(4)	(5)	(6)
Hotspot	0.028	0.016	0.068	0.056	0.102	0.087
	(0.52)	(0.30)	(1.28)	(1.04)	(1.39)	(1.31)
Hotspot×COVID	-0.116	-0.361***	-0.186***	-0.391***	-0.598***	-0.558***
	(-1.66)	(-8.06)	(-2.95)	(-7.12)	(-8.87)	(-8.06)
Hotspot×COVID×{Perf}	0.332*	0.250	0.621***	0.634***	1.392***	1.344***
	(2.07)	(1.41)	(4.42)	(4.06)	(10.06)	(10.25)
HiSCI	0.057	0.066	0.064	0.053	0.224**	0.151
	(0.60)	(0.80)	(0.73)	(0.69)	(2.13)	(1.62)
HiSCI×COVID	-0.088	-0.665***	-0.032	-0.498***	-0.636***	-0.542***
	(-0.53)	(-6.84)	(-0.23)	(-5.76)	(-10.41)	(-6.11)
HiSCI×COVID×{Perf}	0.354*	0.833***	0.297*	0.719***	0.843***	0.690***
	(1.97)	(7.87)	(1.83)	(5.63)	(6.23)	(5.25)
GEO	0.098	0.110	0.094	0.130	0.015	0.081
	(0.49)	(0.56)	(0.43)	(0.58)	(0.09)	(0.47)
GEO×COVID	-0.165	-0.300	-0.277	-0.726***	-0.584***	-0.476***
	(-1.27)	(-1.44)	(-1.46)	(-3.15)	(-5.25)	(-2.95)
GEO×COVID×{Perf}	-0.335	-0.757***	0.038	0.565**	0.308**	0.207
	(-1.41)	(-3.48)	(0.15)	(2.28)	(2.16)	(1.14)
{Perf}	0.698	0.653	0.378	0.376	0.156	0.077
	(1.35)	(1.12)	(1.49)	(1.24)	(1.24)	(0.54)
COVID×{Perf}	1.045***	1.552**	0.243	0.043	-0.768***	-0.630***
	(3.00)	(2.91)	(1.36)	(0.17)	(-12.45)	(-6.91)

Table 2.6: Fund Future Return Regressions, Conditioning on Fund Manager Skill (Continued)

	Perf_CAPM		*Perf_Car4*		*Perf_BB*	
	(1)	(2)	(3)	(4)	(5)	(6)
Hotspot×{Perf}	0.217	0.222	0.167	0.161	0.049	-0.014
	(1.34)	(1.43)	(1.07)	(1.09)	(0.42)	(-0.14)
HiSCI×{Perf}	0.193*	0.123	0.249**	0.187	-0.016	-0.030
	(2.10)	(1.33)	(2.26)	(1.64)	(-0.12)	(-0.24)
GEO×{Perf}	-0.160	-0.108	-0.118	-0.116	-0.043	-0.031
	(-0.75)	(-0.50)	(-0.49)	(-0.48)	(-0.24)	(-0.17)
Flow		0.001		0.002		0.003
		(0.58)		(0.89)		(1.23)
HotspotCorp		0.017***		0.020***		0.021***
		(2.98)		(3.95)		(3.98)
LnAge		0.217**		0.200**		0.051
		(2.20)		(2.38)		(0.99)
LnAUM		-0.010		0.005		0.025
		(-0.20)		(0.11)		(0.77)
Fund_Mgt_Fees		-0.080		-0.078		-0.032
		(-0.87)		(-0.87)		(-1.30)
Fund_Expense		-3.667		-5.929		-9.294
		(-0.18)		(-0.29)		(-0.50)
Fund_Turnover		0.020		0.019		0.010
		(0.36)		(0.35)		(0.19)
$VRet_{q-1}$		1.761		2.234		1.909
		(0.20)		(0.25)		(0.23)
Observations	46,494	43,297	46,494	43,297	53,458	49,723
Adjusted R^2	0.796	0.801	0.794	0.800	0.794	0.801
Fund Style × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

× *{Perf}* is also positive in most cases (-11.1 bp, 24.3 bp, 78.7 bp; *t*-values: -0.70, 1.54, 6.36), suggesting that high skilled manager typically outperform the benchmark group (unconnected managers) even when in the midst of the COVID disaster. Therefore, skilled managers were able to adjust or ignore the salient negative information and position their portfolios to earn normal returns. The results are similar for fund managers in hotspots (*Hotspot* × *COVID*). Thus, *HiSCI* unskilled managers were negatively affected by the salience bias, but high skilled managers were not.

These results are not due to momentum. First, columns 3 and 4 calculate alpha using the Carhart-4 factors, which includes the momentum factor. These results show similar or stronger results than the CAPM results in columns 1 and 2 that do not account for the momentum risk factor. Second, mutual fund investors may chase returns and move into funds with high returns in the past. To control for this, we control explicitly for funds' *Flow* and $Vret_{q-1}$ in columns 2, 4, and 6. Regardless of whether we include these controls or not, the results remain similar (and even strengthen in some cases).

Overall, these results are in line with hypotheses H3a and H3b. Low skilled managers appear to be badly affected by the salience of being in a COVID hotspot or being connected to one of those hotspots. These connected low skilled managers sell more stock and earn inferior returns relative to high skilled or unconnected managers. On the other hand, skilled managers do not suffer from this negative salience bias and earn returns similar to funds that are unconnected and not in hotspots. However, there is little evidence that skilled or unskilled managers benefit from the social connections.

2.4.4 Robustness

HiSCI Funds Far from Hotspots

Geographic proximity to COVID hotspots may contaminate the results as fund managers that work close to hotspots may be affected in a similar way as fund managers in hotspots. The earlier tables control for this using *GEO*, which controls for the distance from the fund and the nearest hotspot. However, to further ensure that the results are not driven by funds close to hotspots, we restrict *HiSCI* funds to only funds that are located at least 100 miles distant from a hotspot. We then rerun Tables 2.3 and 2.6 with this altered definition of *HiSCI*, with results reported in Tables IA.1 and IA.2, respectively.

Our results remain robust to this altered definition of *HiSCI* funds. Tables IA.1 and IA.2 have similar results to their predecessors with funds in the hotspots continuing to sell more shares and see lower returns concentrate among lower skilled managers. Remarkably, the results still exhibit significant social connectedness effects, with *HiSCI* managers located far from the hotspots also selling more shares and seeing lower returns concentrate among lower skilled managers. Consequently, we conclude that geographic proximity to hotspots is not driving the *HiSCI* effect.

Cross-Sectional Analysis

We also conduct cross-sectional analysis on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties in 2020Q1. We use the parsimonious specification below:

$$\Delta H_{i,j,q} = \alpha + \beta_1 \, Hotspot + \gamma_1 \, HiSCI + \theta_1 \, GEO$$

$$+\vartheta_i + \pi_j + \tau_q + \delta_{j,Ind} + \varepsilon_{i,j,q}, \tag{2.3}$$

Table IA. 1: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q1, HiSCI Funds Far from Hotspots

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between 2016Q1 and 2020Q1. *HiSCI* counties are defined as non-hotspot counties whose aggregate SCI to hotspot counties are in the top quartile and are restricted to counties at least 100 miles away from a hotspot. *Share Change (%)* is the change in the number of shares scaled by prior quarter number of shares. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change (%)*	(2) *Share Change (%)*	(3) *Share Change (%)*	(4) *Share Change (%)*	(5) *Share Change (%)*	(6) *Share Change (%)*
Hotspot	-1.145	-1.745	0.571	0.292		
	(-1.11)	(-1.41)	(0.53)	(0.27)		
Hotspot×COVID	-5.483**	-9.204***	-8.555***	-8.118***	-8.462***	-8.102***
	(-2.10)	(-3.57)	(-3.77)	(-3.64)	(-4.03)	(-3.95)
HiSCI	-0.530	-1.130	-2.994***	-3.501***		
	(-0.48)	(-0.87)	(-3.10)	(-3.59)		
HiSCI×COVID	-5.016*	-8.737***	-7.911***	-7.409***	-9.296***	-8.911***
	(-1.89)	(-3.33)	(-3.27)	(-3.14)	(-3.52)	(-3.52)
GEO		-1.858	-0.795	-1.302		
		(-1.47)	(-0.74)	(-1.16)		
GEO×COVID		-11.933***	-9.856**	-9.305**	-11.141***	-10.707***
		(-2.72)	(-2.47)	(-2.34)	(-2.79)	(-2.70)
COVID	-1.802	1.919	2.611		1.827	
	(-0.98)	(1.07)	(1.47)		(1.10)	
Observations	5,023,657	5,023,657	4,876,486	4,876,295	4,836,605	4,836,418
Adjusted R^2	0.001	0.001	0.037	0.050	0.059	0.074
Controls	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	Yes	No	Yes
Fund FE	No	No	No	No	Yes	Yes
Quarter FE	No	No	No	Yes	No	Yes
Fund × Industry FE	No	No	No	No	Yes	Yes

Table IA. 2: Fund Future Quarterly Return Regressions, Conditioning on Fund Manager Skill, HiSCI Funds Far from Hotspots

This table reports panel regressions of future quarterly fund returns between 2016Q1 and 2020Q1, conditioning on fund manager's CAPM alpha in the past 5 years (col 1-2), Carhart 4 alpha in the past 5 years (col 3-4), and fund manager skill calculated as per Berk and van Binsbergen (2015) (col 5-6). *HiSCI* counties are defined as non-hotspot counties whose aggregate SCI to hotspot counties are in the top quartile and are restricted to counties at least 100 miles away from a hotspot. The dependent variable is the value-weighted fund return less fees in quarter q+1 ($VRet_{q+1}$). *Perf_CAPM/Perf_Car4/Perf_BB* is an indicator variable set to 1 if the fund is in the top-tercile of CAPM alpha, Carhart-4 alpha or Berk and van Binsbergen (2015) manager skill and 0 otherwise. Control variables are identical to those in Table 2. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and quarter, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

Skill Measure *{Perf}*	*Perf_CAPM*		*Perf_Car4*		*Perf_BB*	
	(1)	(2)	(3)	(4)	(5)	(6)
Hotspot	0.057	0.044	0.098	0.083	0.133	0.116
	(1.16)	(0.95)	(1.53)	(1.25)	(1.59)	(1.52)
Hotspot×COVID	-0.120	-0.447***	-0.200**	-0.454***	-0.672***	-0.668***
	(-1.52)	(-8.81)	(-2.80)	(-5.83)	(-8.76)	(-8.04)
Hotspot×COVID×{Perf}	0.287*	0.238	0.598***	0.581***	1.361***	1.353***
	(1.93)	(1.63)	(4.53)	(4.44)	(9.94)	(10.50)
HiSCI	0.148	0.153	0.155	0.135	0.310**	0.232**
	(1.41)	(1.37)	(1.36)	(1.11)	(2.70)	(2.24)
HiSCI×COVID	-0.101	-0.938***	-0.072	-0.689***	-0.842***	-0.851***
	(-0.48)	(-8.74)	(-0.43)	(-6.04)	(-13.26)	(-7.35)
HiSCI×COVID×{Perf}	0.237	0.857***	0.237	0.582***	0.708***	0.671***
	(1.19)	(5.57)	(1.40)	(3.49)	(4.24)	(4.66)
GEO	0.156	0.172	0.157	0.185	0.164	0.191
	(0.79)	(0.82)	(0.68)	(0.75)	(0.98)	(1.07)
GEO×COVID	-0.213	-0.718***	-0.307	-1.041***	-0.995***	-0.873***
	(-1.28)	(-3.45)	(-1.26)	(-4.24)	(-10.09)	(-5.19)
GEO×COVID×{Perf}	-0.171	-0.332	0.190	0.894***	0.726***	0.586***
	(-0.81)	(-1.41)	(0.66)	(3.42)	(5.13)	(3.37)
Observations	46,494	43,297	46,494	43,297	53,458	49,723
Adjusted R^2	0.796	0.801	0.794	0.800	0.794	0.801
Other Perform Interactions	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Fund Style × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.3 shows that funds located in hotspot counties and high-SCI counties sell more share holdings even when only look at the one quarter of sample.

Alternate Definitions of Hotspots

To check whether our results are robust to the definition of COVID-19 hotspots, we use 2,500 cases and 1.5 cases per 100,000 population as the threshold for *Hotspot* county classification instead of the 2,000 threshold. In Panel A and B of Table IA.4, we replicate the results on cross-sectional and panel analysis except adjusting the definition of *Hotspot* using the alternative definitions above. Regardless of the metric we used, the results for *Hotspot* and *HiSCI* remain with both these variables being negatively related to fund stock holdings.

Alternate Definitions of Selling

The metric in Table 3, *Share Change (%)*, does not account for any new positions the fund takes as the % change will return an undefined value if we divide the change in shares by 0. To avoid this problem, we develop an alternate metric, *Share Change Per Shrout*, which is the change in number of shares scaled by the total number of shares outstanding of the firm. This avoids the divide by 0 issue and allows us to include new positions.

We replicate Table 2.3 using *Share Change Per Shrout* and present the results in Table IA.5. The *HiSCI* × *COVID* coefficients remain significantly negative regardless of what kind of fixed effects we use. Even though the *Hotspot* × *COVID* results lose significance, with an alternate measure of share change, being socially connected or in a hotspot is associated with a reduction in share ownership.

Table IA. 3:Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2020Q1 (COVID Outbreak Period Only)

This table reports the results of the cross-sectional analysis on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties in 2020Q1. The dependent variables are reported on top of each column. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *Share Change Per Shrout* is the change in share holdings scaled by the total number of shares outstanding (in 10^{-5}s), which is calculated as the change in share number multiplied by the stock price at the beginning of the quarter. Refer to Appendix A for detailed variable definitions of other variables. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and **** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Share Change (%)	*Share Change (%)*	*Share Change (%)*	*Share Change Per Shrout*	*Share Change Per Shrout*	*Share Change Per Shrout*
Hotspot	-9.580***	-7.777***	-3.920***	-3.666**	-2.393	-2.680*
	(-3.58)	(-3.98)	(-2.67)	(-2.44)	(-1.50)	(-1.69)
HiSCI	-13.642***	-10.491***	-8.065***	-5.720***	-4.534***	-4.478**
	(-4.72)	(-4.35)	(-3.59)	(-3.31)	(-2.60)	(-2.54)
GEO		-4.658	-2.680		-0.322	-0.570
		(-1.33)	(-0.83)		(-0.16)	(-0.29)
Flow		1.572***	1.439***		0.598***	0.627***
		(12.02)	(10.76)		(6.65)	(6.88)
HotspotCorp			-0.329			-0.496
			(-0.45)			(-1.00)
LnAge			1.168			-4.258***
			(0.54)			(-4.51)
LnAUM			-4.664***			1.535*
			(-3.82)			(1.72)
Fund_Mgt_Fees			-459.985			66.898
			(-1.40)			(0.33)
Fund_Expense			-1.539			0.159
			(-1.55)			(0.37)
Fund_Turnover			-0.003			-0.373
			(-0.01)			(-1.45)
Observations	347,168	330,641	328,810	372,613	353,727	351,710
Adjusted R^2	0.048	0.098	0.108	0.222	0.182	0.183
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes

Table IA. 4: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors: Alternative Definitions of Hotspot

This table provides robustness results of the analysis on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties using alternative definition of hotspot counties. Instead of using 2,000 as the threshold number of COVID cases by March 30, 2020, we use 2,500 cases and 1.5 cases per 100,000 population, respectively, to define *Hotspot*. Panel A reports cross sectional regression results using the 2020Q1 data, and Panel B reports results for the full sample between 2016Q1and 2020Q1. The dependent variables are *Share Change (%)* which is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *HiSCI* is defined as non-hotspot counties whose aggregate SCI to hotspot counties are in the top quartile. *GEO* is defined as non-*Hotspot* counties, which are within 100 miles of any one of the hotspot counties. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Alternative Definition of Hotspot County in cross sectional analysis of *Share Change %*: 2016Q1(COVID Period)

	No. Cases >= 2,500		No. Cases per 100,000 Population >= 1.5	
	(1)	(2)	(3)	(4)
Hotspot	-7.148**	-1.872	-9.951***	-3.930***
	(-2.41)	(-1.15)	(-3.75)	(-2.70)
HiSCI	-11.550***	-10.410**	-13.790***	-7.874***
	(-4.09)	(-2.07)	(-4.77)	(-3.14)
GEO		4.142		-1.749
		(0.80)		(-0.45)
Observations	347,168	328,810	347,168	328,810
Adjusted R^2	0.045	0.108	0.048	0.108
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes

Panel B: Alternative Definition of Hotspot County in panel regression of *Share Change %*: 2016Q1 to 2020Q1

	No. Cases >= 2,500		No. Cases per 100,000 population >= 1.5	
	(1)	(2)	(3)	(4)
Hotspot	0.453		1.056	
	(0.40)		(1.04)	
Hotspot × *COVID*	-5.917**	-5.841***	-8.985***	-8.665***
	(-2.51)	(-2.73)	(-4.10)	(-4.38)
HiSCI	0.325		-1.563*	
	(0.32)		(-1.94)	
HiSCI × *COVID*	-12.551**	-14.946***	-9.620***	-10.794***
	(-2.33)	(-2.69)	(-3.35)	(-3.68)
GEO	-1.849*		0.847	
	(-1.66)		(0.95)	
GEO × *COVID*	3.647	4.645	-5.096	-5.130
	(0.67)	(0.84)	(-1.17)	(-1.16)
COVID	0.622		2.966*	
	(0.39)		(1.72)	
Observations	4,876,486	4,836,418	4,876,486	4,836,418
Adjusted R^2	0.036	0.074	0.037	0.074
Controls	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Quarter FE	No	Yes	No	Yes
Fund × Industry FE	No	Yes	No	Yes

Table IA. 5: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors: Alternate Measurement of Share Change

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between Q1 of 2016 and Q1 of 2020. *Share Change Per Shrout* is the change in share holdings scaled by the total number of shares outstanding (in 10^{-5}s). *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Controls are similar to Table 2. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change Per Shrout*	(2) *Share Change Per Shrout*
Hotspot	-1.537	
	(-1.59)	
Hotspot×COVID	-0.699	0.791
	(-0.29)	(0.25)
HiSCI	-1.391	
	(-1.39)	
HiSCI×COVID	-4.720*	-8.633**
	(-1.65)	(-2.31)
GEO	-0.830	
	(-0.85)	
GEO×COVID	1.632	4.907
	(0.64)	(1.48)
COVID	-3.535*	
	(-1.75)	
Observations	5,335,611	5,309,850
Adjusted R^2	0.000	0.018
Controls	Yes	Yes
Firm FE	No	Yes
Fund FE	No	Yes
Quarter FE	No	Yes
Fund × Industry FE	No	Yes

Length of Panic Period

We have shown in Table 2.3 that being in or socially connected to COVID hotspots was associated with intensified fund stock selling during 2020Q1. In Table IA.6, we examine whether this association continued after the COVID outbreak quarter by using in indicator *PostCOVID* which equals 1 if the time period is 2020Q2 and 0 otherwise. The results indicate much weakened effects of *Hotspot* × *PostCOVID* and *HiSCI* × *PostCOVID* with signs flipping between the share change measures used, suggesting most of the panic-driven stock selling occurred during the outbreak quarter of 2020Q1.

Firm and Fund Characteristics

One concern is that the selling during the pandemic may be driven by firm characteristics. To control for this, Table 2.3 already includes firm fixed effects to account for time-invariant unique firm characteristics. Furthermore, to ensure that firm characteristics are not driving our results, in Table IA.7 we include a host of firm controls comprising: firm size, analyst coverage, return skewness (maximum daily return in the previous quarter), institutional/retail ownership, location of the firm in a hotspot, idiosyncratic volatility, exposure to economic uncertainty (uncertainty beta), and stock price. The results remain qualitatively similar even with these extra controls and with the reduction of sample size.

Furthermore, we also specifically investigate whether the selling results in Table 2.3 are driven by specific firm and fund characteristics that have been identified in the social finance literature.[13] For example, Hirshleifer et al. (2020) identifies size, investor attention, retail

[13] It is worth noting that fund return regressions cannot include firm-level characteristic controls as funds have diverse portfolios covering hundreds of different firms.

Table IA. 6: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q2: Focusing on the Post-COVID Period

This table reports the results of the panel regression on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between Q1 of 2016 and Q1 of 2020. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *Share Change* is the change in share number (in thousands) after adjusting for stock splits. *Share Change Per Shrout* is the change in share holdings scaled by the total number of shares outstanding (in 10^{-5}s). *PostCOVID* is an indicator variable set to 1 if the time period is 2020Q2 and 0 otherwise. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change (%)*	(2) *Share Change (%)*	(3) *Share Change Per Shrout*	(4) *Share Change Per Shrout*
Hotspot	0.598		-0.745	
	(0.59)		(-0.57)	
Hotspot×PostCOVID	0.695	1.125	33.526	3.127
	(0.27)	(0.44)	(0.86)	(0.96)
HiSCI	-1.464*		-1.105	
	(-1.93)		(-1.01)	
HiSCI×PostCOVID	3.845**	2.881	-1.797	-2.529
	(2.02)	(1.51)	(-0.28)	(-0.76)
GEO	-0.104		-0.613	
	(-0.13)		(-0.61)	
GEO×PostCOVID	4.927*	5.278**	-2.037	4.234
	(1.89)	(2.03)	(-0.36)	(1.39)
PostCOVID	-8.773***		12.244	
	(-4.72)		(1.24)	
Flow	0.941***	0.959***	0.420***	0.358***
	(21.23)	(19.34)	(10.28)	(8.60)
HotspotCorp	0.028	-0.416*	-0.263**	-0.055
	(0.56)	(-1.80)	(-2.02)	(-0.98)
LnAUM	-0.981*	-4.675***	-0.451	16.398***
	(-1.84)	(-2.64)	(-0.67)	(5.81)
LnAge	1.168***	-1.057	0.067	-0.436
	(4.29)	(-0.81)	(0.26)	(-0.53)
Fund_Mgt_Fees	-0.786***	-0.313	0.738*	-0.526
	(-3.23)	(-1.33)	(1.72)	(-0.47)
Fund_Expense	-3.149**	5.060	-3.322	8.874
	(-2.03)	(1.01)	(-1.01)	(1.13)
Fund_Turnover	-0.428	-0.015	0.075	-1.100*
	(-1.38)	(-0.06)	(0.26)	(-1.81)

$VRet_{q-1}$	-0.058**	-0.230***	0.360***	0.359***
	(-2.22)	(-4.20)	(3.37)	(3.33)
Observations	5,206,634	5,165,169	5,693,653	5,668,621
Adjusted R^2	0.036	0.072	0.000	0.420
Firm FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Quarter FE	No	Yes	No	Yes
Fund × Industry FE	No	Yes	No	Yes

Table IA. 7: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q2: Additional Firm Characteristic Controls

This table reports the results of the panel regression on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between 2016Q1 and 2020Q1. It is similar to Table 3, column 6. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *Share Change Per Shrout* is the change in share holdings scaled by the total number of shares outstanding (in 10^{-5}s). *Size* is the natural log of market capitalization of the firm. *NumEst* is the analysts coverage providing the estimates of earnings. *IO* is the percent of total institutional ownership. *HOTLOC* is indicator variable equals one if the firm's headquarter is located at hotspot county. MAX is the daily maximum return in previous month. *IVOL* is the standardized idiosyncratic volatility based on the Fama-French 3 factor model. *Unc_Beta* measures firms' exposure to economic uncertainty (Baker, Bloom, and Davis 2016). *PRC* is the level of stock price. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change (%)*	(2) *Share Change (%)*
Hotspot	1.194	
	(1.14)	
Hotspot×COVID	-9.537***	-9.688***
	(-4.26)	(-4.80)
HiSCI	-1.261	
	(-1.52)	
HiSCI×COVID	-10.781***	-12.883***
	(-3.79)	(-4.42)
GEO	0.072	
	(0.08)	
GEO×COVID	-5.464	-5.662
	(-1.38)	(-1.41)
COVID	5.277***	
	(3.00)	
Flow	0.959***	0.981***
	(20.16)	(18.50)
HotspotCorp	0.008	-0.457*
	(0.13)	(-1.79)
LnAge	-1.344**	-1.625
	(-2.56)	(-0.89)
LnAUM	1.271***	-1.541
	(4.88)	(-1.04)
Fund_Mgt_Fees	-0.792***	-0.459
	(-2.78)	(-1.47)
Fund_Expense	-3.045*	2.862
	(-1.86)	(0.60)
Fund_Turnover	-0.152	-0.032
	(-0.43)	(-0.11)
$VRet_{q-1}$	0.007	-0.120*
	(0.26)	(-1.73)
Size	0.319**	4.894***

	(2.13)	(9.90)
NumEst	-0.036**	-0.043
	(-2.34)	(-1.09)
IO	2.302***	1.586***
	(4.00)	(2.64)
HOTLOC	0.091	1.601
	(0.38)	(0.56)
MAX	-0.169***	-0.142***
	(-7.20)	(-6.70)
IVOL	8.335**	-18.365**
	(2.43)	(-2.38)
Unc_Beta	-1.974	-0.196
	(-1.33)	(-0.08)
PRC	-0.000*	0.000***
	(-1.92)	(2.75)
Observations	4,291,099	4,253,723
Adjusted R^2	0.040	0.071
Firm FE	No	Yes
Fund FE	No	Yes
Quarter FE	No	Yes
Fund × Industry FE	No	Yes

ownership, and glamour as affecting the length of social networks' impact on earnings announcement returns. Retail ownership is also a proxy for the prevalence of noise traders (De Long et al. 1990). Consequently, we investigate these factors by examining interactions with size, analyst coverage, institutional ownership (the inverse of retail ownership), and book-to-market. Further, a firm headquartered in a COVID hotspot may be more economically affected by the pandemic than firms in other regions. We consequently investigate if our results relate to the firm's location in a hotspot.[14]

Table IA.8 Panel A shows the results of the interaction tests with these characteristics. The main effects, *Hotspot* × *COVID* and *HiSCI* × *COVID* remain significantly negative in all cases. The *Hotspot* × *COVID* interaction is only significant for size and analyst coverage, showing that being bigger or having more coverage mitigates some of the selling for *Hotspot* funds. However, institutional ownership, book-to-market, and a firm located in hotspot, has no significant impact on selling when interacted with *Hotspot* during COVID (*Hotspot* × *COVID* × *{Interaction}*). Furthermore, there is no significant impact for any of the variables when interacted with *HiSCI* during COVID (*HiSCI* × *COVID* × *{Interaction}*). Thus, it is apparent that firm characteristics do not play a significant role in our results—the results are driven by social connections and not other firm characteristics.

In addition to firm characteristics, we examine two fund-level characteristics that can influence fund selling. *HotspotCorp*, the amount of the fund's portfolio that is headquartered in hotspots; this is directly related to the likelihood of selling as the fund may wish to divest from

[14] In other robustness tests, we also include an equity mispricing measure based on Rhodes-Kropf et al. (2005) in the stock selling regression and find our results are robust to controlling for mispricing. But since inclusion of this mispricing metric substantially reduces sample size, we do not report results with this variable.

Table IA. 8: Impact of COVID Hotspot and Social Connectedness on Stock Selling of Institutional Investors, 2016Q1 to 2020Q2: Interactions with Firm and Fund Characteristics

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between 2016Q1 and 2020Q1. The regression is similar to Table 2.3 column 6 except that an interaction is included for major firm or fund characteristics. Panel A shows the interactions with firm characteristics comprising: *Large,* set to 1 if the stock's market capitalization is above the median; *Coverage,* set to 1 if the stock's analyst coverage is above the median; *HighIO,* set to 1 if the stock's total institutional ownership is above the median; *HighBM*, set to 1 if the stock's Book-to-Market ratio is above the median; and *HOTLOC*, set to 1 if the firm's headquarter is in hotspot county. Panel B shows the interactions with fund characteristics comprising: *HighHot*, set to 1 if the fund's *HotspotCorp* is above the median; and *HighTurn*, set to 1 if the fund's turnover ratio is above the median. *HighFlow*, set to 1 if the fund's Net Flow is above the median. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Interactions with Firm Characteristics

Characteristic *{Interaction}*	(1) *Large*	(2) *Coverage*	(3) *HighIO*	(4) *HighBM*	(5) *HOTLOC*
Hotspot×COVID	-13.091***	-11.829***	-9.463***	-9.076***	-9.739***
	(-5.24)	(-4.98)	(-4.70)	(-4.58)	(-4.83)
Hotspot×COVID×{Interaction}	6.982***	4.651**	-0.411	-0.772	0.357
	(2.96)	(2.31)	(-0.43)	(-0.62)	(0.29)
HiSCI×COVID	-10.815***	-10.028***	-12.173***	-12.027***	-12.866***
	(-3.54)	(-3.66)	(-4.11)	(-3.96)	(-4.44)
HiSCI×COVID×{Interaction}	-2.275	-4.158	-1.326	-1.525	-0.285
	(-0.60)	(-1.33)	(-1.26)	(-1.05)	(-0.20)
GEO×COVID	-10.375***	-8.362***	-5.295	-5.194	-5.907
	(-3.63)	(-3.10)	(-1.24)	(-1.17)	(-1.49)
GEO×COVID×{Interaction}	8.524*	5.093	-0.630	-2.323*	2.342
	(1.65)	(1.17)	(-0.53)	(-1.73)	(1.61)
COVID×{Interaction}	-4.407**	-3.000*	-0.330	1.917**	1.404*
	(-2.24)	(-1.76)	(-0.51)	(2.38)	(1.66)
Hotspot×{Interaction}	2.652***	1.349**	0.463	-0.447	0.944**
	(2.75)	(1.98)	(1.21)	(-0.83)	(2.20)
HiSCI×{Interaction}	3.793***	1.193*	0.611*	-1.135**	-0.320
	(3.99)	(1.77)	(1.69)	(-2.26)	(-0.73)
GEO×{Interaction}	0.004	0.878	0.074	1.175**	0.077
	(0.00)	(1.23)	(0.17)	(2.07)	(0.17)
{Interaction}	-3.515***	-1.733***	-0.150	0.489	1.121
	(-5.13)	(-3.87)	(-0.42)	(1.08)	(0.40)
Observations	4,253,723	4,253,723	4,253,723	3,495,190	4,253,723
Adjusted R^2	0.071	0.071	0.071	0.070	0.071
Fund Controls	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Fund × Industry FE	Yes	Yes	Yes	Yes	Yes

Panel B: Interactions with Fund Characteristics

Characteristic *{Interaction}*	(1) *HighHot*	(2) *HighTurn*
Hotspot×COVID	-17.173***	-8.525***
	(-5.46)	(-3.91)
Hotspot×COVID×{Interaction}	14.128***	0.795
	(3.73)	(0.21)
HiSCI×COVID	-17.392***	-11.653***
	(-4.00)	(-3.30)
HiSCI×COVID×{Interaction}	9.548*	0.006
	(1.74)	(0.00)
GEO×COVID	-18.186***	4.545
	(-3.81)	(0.61)
GEO×COVID×{Interaction}	22.190***	-14.838*
	(3.07)	(-1.78)
COVID×{Interaction}	-9.756***	-5.637*
	(-3.36)	(-1.82)
Hotspot×{Interaction}	8.592	1.502
	(0.75)	(0.56)
HiSCI×{Interaction}	6.774	0.079
	(0.76)	(0.04)
GEO×{Interaction}	7.626	-2.243
	(1.00)	(-0.89)
{Interaction}	-8.802	0.044
	(-0.81)	(0.05)
Observations	4,253,723	4,253,723
Adjusted R^2	0.072	0.071
Fund Controls	Yes	Yes
Firm Controls	Yes	Yes
Firm FE	Yes	Yes
Fund FE	Yes	Yes
Quarter FE	Yes	Yes
Fund × Industry FE	Yes	Yes

firms that are economically exposed to COVID. *Fund_Turnover*, how much the fund buys/sells over the past 12 months, also is directly related to the likelihood of selling funds; if a fund typically sells more, it will likely sell more during the pandemic.

The results are shown in Table IA.8 Panel B. As can be seen, *Hotspot/HiSCI* × *COVID* × *HighTurn* (a dummy for above median fund turnover) has no significance. Thus, fund turnover has no impact on the selling during the COVID period for funds in hotspots or socially connected to hotspots. Finally, *Hotspot/HiSCI* × *COVID* × *HighHot* has a significantly *positive* impact on funds' stock sales. This is directly counter to the hypothesis noted above; being economically exposed to COVID during the pandemic outbreak should *accelerate* selling, not slow it. Regardless, we conclude that the *Hotspot* × *COVID* and *HiSCI* × *COVID* effects are not being driven by these fund characteristics.

Placebo Test

One concern is that the share sales are simply an industry-wide phenomenon that all fund managers are forced to be part of. To examine this possibility, we test a placebo group that should *not* experience any significant selling: index funds. Index funds do not have active management and do not adjust their holdings according to market conditions. In contrast, the actively management funds in our sample are expected to change stock holdings, particularly during severe market conditions. Consequently, by examining the differences between index funds and active funds in our sample, we distinguish between an industry-wide sale or a reaction limited to only actively managed funds.

Table IA.9 reports the results of this placebo test of index funds. The coefficients for *Hotspot* × *COVID* and *HiSCI* × *COVID* are insignificant in columns 1-2 regardless of the usage of fixed

Table IA. 9: Placebo Test: Impact of COVID-19 and Social Connectedness on Stock Ownership of Institutional Investors, 2016Q1 to 2020Q1, Index Funds Only

This table reports the results of the panel regressions on how the institutional investors' changes in share ownership differ across hotspot, high-SCI, and low-SCI counties between Q1 of 2016 and Q1 of 2020. For this table the sample only includes index funds and excludes any actively managed funds. *Share Change (%)* is the change in the number of shares over the quarter scaled by prior number of shares in percentage form. *COVID* is an indicator variable set to 1 if the time period is 2020Q1 and 0 otherwise. Controls are similar to Table 2. Refer to Appendix A for detailed variable definitions. Standard errors are double clustered by fund and firm, with *t*-statistics reported in parentheses. *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

	(1) *Share Change (%)*	(2) *Share Change (%)*
Hotspot	2.370**	
	(2.50)	
Hotspot×COVID	3.055	3.974
	(0.99)	(1.25)
HiSCI	3.151**	
	(2.54)	
HiSCI×COVID	-2.989	-3.415
	(-0.81)	(-0.93)
GEO	-0.526	
	(-0.57)	
GEO×COVID	2.795	1.985
	(0.99)	(0.68)
COVID	-6.420**	
	(-2.42)	
Observations	5,249,883	5,238,337
Adjusted R^2	0.080	0.151
Controls	Yes	Yes
Firm FE	No	Yes
Fund FE	No	Yes
Quarter FE	No	Yes
Fund × Industry FE	No	Yes

effects. There does not appear to be a general selling of stocks in index funds during the COVID period. We thus conclude that our results are not due to a general selloff, but due to choices by active managers who have reacted to the COVID crisis.

2. 5 Conclusion

There is a debate in the literature on whether informal social connections promote information sharing or intensify biased heuristics that causes poor trading decisions. Previous theoretical and empirical work has shown negative consequences of social interactions on investment decisions for retail investors and this paper extends the work into the realm of professional money managers.

This paper shows that COVID-19 intensified mutual fund stock selling in counties with high levels of the disease during 2020Q1. Further, fund managers socially connected to these hotspots also sold more stock during this period. Finally, the results indicate that social connections made low skilled fund managers more vulnerable to salience bias during this episode; these fund managers acted on the information from social connections but used the information to make poor trades due to the salient fear of COVID-19. On the other hand, skillful managers appeared to have been able to ignore the salience bias from being directly exposed to or socially connected to high levels of COVID. Therefore, our evidence suggests that the impact of social media on investing depends on fund manager skill.

Even though our results are specific to this pandemic outbreak period, owing to the unprecedented economic impact and the exogenous nature of the shock, the social connection effects we document are highly economically significant and largely free of endogeneity. Our findings that social networks transmitted both salience bias and valuable information among institutional investors under stress, and that the effects depended on manager skill, suggest that the impact of social connectivity on investor behavior warrants future research.

Our work has important practical and policy implications. If social interactions intensify behavioral biases among many professional money managers during times of crisis, social connectivity can instigate market volatility and destabilize the financial markets. Therefore, it pays central banks and policy makers to be wary of the downside of social connectedness and design mechanisms to stem investor irrationality through social networks.

2. 6 Appendix

Appendix 2.A Variable Definitions

Variable	Definition
AUM	Total equity assets under management of a mutual fund, calculated using the equity assets in the CRSP universe.
COVID	Indicator variable set to 1 if the time period is 2020Q1; 0 otherwise.
Epic	Indicator for stocks in the 10 or 24 industries with the lowest equal-weighted returns during 2020Q1.
Flow	Net flow into the fund scaled by previous quarter *AUM*. $$Flow_{i,t} = \frac{AUM_{i,t} - AUM_{i,t-1} \times \left(1 + R_{i,t}\right)}{AUM_{i,t-1}}$$
Fund_Expense	Fund expense ratio. Ratio of total investment that shareholders pay for the fund's operating expenses. Fund expense may include waivers and reimbursements, causing it to appear less than the fund management fee.
Fund_Mgt_Fees	The fee is calculated using ratios based on the line items reported in the Statement of Operations. The management fee can be offset by fee waivers and/or reimbursements which will make this value differ from the contractual fees found in the prospectus. Reimbursements can lead to negative management fees.
Fund_Turnover	Fund turnover ratio. Minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month Total Net Assets of the fund.
GEO	Indicator for non-hotspot counties that are geographically close to hotspots, defined as within 100 miles of any one of the hotspot counties.
HotspotCorp	Percentage of a fund's *AUM* that is invested in firms whose headquarters are located in a *Hotspot* county.
HiSCI	Indicator variable for a non-hotspot county whose aggregate *SCI* to *Hotspot* counties is in the top quartile.
LnAge	The natural logarithms of 1 + the Fund's age in years.
Hotspot	Indicator for counties which have more than 2,000 COVID-19 cases reported by March 30, 2020.
Perf_BB	Indicator variable set to 1 if a fund is in the top tercile of value added as defined by Berk and van Binsbergen (2015). Fund manager performance measurement begins in 2010 and ends 1 year before the date being tested (i.e., end date of 2018 for any 2019 mutual fund regressions).
Perf_CAPM	Indicator variable for high fund manager skill defined similarly to *Perform* with historical alpha estimated using the CAPM model instead of Carhart 4 factors.

Perf_Car4	Indicator variable set to 1 if a fund is in the top tercile of alpha in the previous period. The alpha is measured using returns adjusted for the Carhart (1997) 4 factors over the rolling past 5 years using monthly data.
SCI	Social Connectedness Index from Facebook. See Appendix A.2 on how it is developed.
Share Change (%)	The scaled change in number of shares held by a mutual fund, defined as the change in number of shares over the quarter scaled by number of shares in the previous quarter, in percentage form.
Share Change Per Shrout	The scaled change in number of shares held by a mutual fund, defined as the change in number of shares over the quarter scaled by number of shares outstanding (in 10^{-5}s).
VRet	Fund return (net of cost such as management fees and expenses) in percentages at the end of the current quarter which is calculated as the value weighted return of each fund class, using *AUM* as the weight.

Appendix 2.B Facebook Social Connectedness Index (SCI) Methodology

This methodology is excerpted from Facebook's Social Connectedness (*SC*) data site available at https://www.facebook.com/help/geoinsights/880664345608937/?helpref=hc_fnav&bc[0]=SPACO%20Help%20Center&bc[1]=Disease%20Prevention%20Maps

> We use an anonymized snapshot of all active Facebook users and their friendship networks to measure the intensity of connectedness between locations. Locations are assigned to users based on their information and activity on Facebook, including the stated city on their Facebook profile, and device and connection information. Our primary measure of Social Connectedness, SC, between two locations *i* and *j* is:
>
> $SC_{,ij}$ = $FB\_Connections_{i,j}$ / (FB_Users_i * FB_Users_j)
>
> Here, FB_Users *i* and FB_Users *j* are the number of Facebook users in locations *i* and *j*, and $FB\_Connections_{i,j}$ is the number of Facebook friendship connections between the two.
>
> $SC_{i,j}$, therefore, measures the relative probability of a Facebook friendship link between a given Facebook user in location *i* and a user in location *j*. Put differently, if this measure is twice as large, a Facebook user in *i* is about twice as likely to be connected with a given Facebook user in *j*.
>
> In each dataset, we scale the measure to have a fixed max value (by dividing the original measure by the max and multiplying by 1,000,000,000) and the lowest possible value of 1. Since absolute values of $SC_{i,j}$ are not meaningful, this re-scaling does not affect the interpretation. We also round the measure to the nearest integer.

The high-SCI indicator, *HiSCI*, is calculated in two steps: In the first step, the *SCI* score of county *i* is calculated as the average SC between county *i* and all hotspot counties, where *h* is the number of *Hotspot* counties:

$$SCI_i \ = \sum_1^h SC_{ij} \,/\, h$$

In the second step, the *SCI* scores are sorted from highest to lowest and the top quartile of non-hotspot counties are flagged as *HiSCI* counties. All remaining counties are flagged as *HiSCI* = 0 and *LoSCI* = 1.

Chapter 3 Whose Attention Matters? Evidence from Media News Sentiment

3. 1 Introduction

Can paying attention to *who* is paying attention help us anticipate market reactions to news? We know that a large literature[15] provides empirical evidence that the release of media news is often accompanied with mixed anomalous return responses such as return drift or reversal, interpreted as underreaction and overreaction. We also know that a growing literature[16] in finance shows that investor (in)attention is associated with anomalous return patterns. But the evidence on the impact of news and investor attention on markets is mixed and conflicting. For example, some studies (Cohen and Frazzini 2008, en-Rephael, Da, and Israelson, 2017; Ben-Rephael, Da, Easton, and Israelson, 2018) show that institutional attention plays a stabilizing role on financial markets, while others, such as Ma, Xiong, and Feng (2020) document a destabilizing role for institutional attention on news releases, at least in emerging markets. Barber and Odean, (2008) and Da, Engelberg, and Gao, (2011) document that retail attention destabilizes financial markets but recent studies including Liu, Peng, and Tang, (2019) and Zhou (2020) find evidence for a stabilizing role played by retail attention. In this study, we seek to reconcile the evidence on return responses to news by identifying the different types of investors paying attention, retail (local and national) versus institutional, while controlling for news sentiment and news complexity. Although recent papers have explored each of these separately or in small subsets, no work to our knowledge has

[15] Pritamani and Singal (2001); Tetlock (2007); Savor (2012); Tetlock (2014)
[16] Hirshleifer and Teo (2003), Hirshleifer, Hou, Teoh, and Zhang (2004), Hou (2007), Peng and Xiong (2008), Cohen and Frazzini (2008), Barber and Odean (2008), DellaVigna and Pollet, 2009; Hirshleifer, Lim, and Teoh (2011), Da, Engelberg, and Gao (2011)

endeavoured to pull apart and understand the data as we do. This leads to a nuanced understanding of the impact of news and attention on financial markets.

The market response to news may vary with the type of investor paying attention due to the different roles played by institutional and retail investors. On the one hand, recent studies show that institutional investors' attention stabilizes market prices and lack of this attention leads to price underreaction to news. For example, Ben-Rephael, Da and Israelson (2017) find that the post announcement return drift weakens conditioning on institutional abnormal attention. On the other hand, a strand of literature finds that retail investors' attention is a source of price instability, leading to price overreaction and reversal. For instance, Da, Engelberg and Gao (2011) find that abnormal retail attention will first induce higher stock prices in 2 weeks and an eventual price reversal within the year. In sum, it suggests that institutional inattention could induce return drift and retail attention could induce return reversal, which could, at least partially, explain anomalous return responses to news. Studies that look for market reaction to news will provide unreliable conclusions if the type of investor paying attention is not controlled for, a concern that is not widely recognised. Common proxies like trading volume, previous returns, or news profiling the firm in question are indirect measures of attention and are unable to identify its' source (or indeed even its' existence). Our study shows that there are news stories that attract only retail investor attention, or only institutional investor attention, or both or neither. We determine that these situations have different implications for market price reaction to the news, itself a new finding.

Utilizing mass media news sources and textual analysis, we identify the sentiment of the news and the complexity of the news, and we identify attention to the news from institutional and retail investors using Bloomberg institutional investor attention indices and Google trends data. We use these data to investigate the interaction of news sentiment (positive, negative or neutral), news type (straightforward or complex) and investor-type attention (retail or institutional). Specifically, we use the Topic-Adaptive Syntax (TASA) Approach proposed in Babolmorad and Massoud (2020) to identify the sentiment of media news and categorize news into positive, negative, and neutral news. In addition, TASA approach is able to help us categorize news into financial-related news which is straightforward to interpret and business-related news which is complex to interpret.

We document the following primary results. First, we find evidence of price reversal when retail investors pay attention and institutional attention is absent, but no reversal when institutional attention is high, linking results from several strands of the literature. Different from Da, Engelberg and Gao (2001), who document the price reversal without specifying events, we contribute to the literature by looking at the market response to a specific type of events - media news releases. In addition, we use textual analysis to determine the valence of the sentiment of media news, which is different from Frank and Sanati (2018) who document price reversal by identifying the news sentiment using return response on the media news day, with the assumption that the response fully reflects the news and only the news.

When focusing on the sentiment of news, we find that the return reversal induced by retail abnormal attention is associated with positive news, not negative. The fact that the reversal comes with positive news rather than negative is consistent with constraints for short-selling that face even institutional investors (see, for instance, Gervais, Kaniel, and Mingelgrin 2001). This finding suggests that the interaction of news sentiment and the type of investor attention determines the market response following the release of news.

Second, we observe return drift on positive news when there is a lack of institutional attention. In other words, the return drift gets weaker or even disappears when there is institutional attention. Ben-Rephael, Da and Israelson (2017) find that the post announcement return drift on earnings or recommendation changes weakens or disappears conditioning on institutional attention, but do not consider news more broadly, as we do. Cohen and Frazzini (2008) find return drift in price if institutional investors are inattentive, exploiting large return movements as proxies for news rather than identifying news events and measuring news sentiment as we do. Our refinement of investor-type attention and news sentiment leads us to a more nuanced understanding of the impact of inattention. Rather than a blanket result, that a lack of attention results in a drift in prices, we find that inattention leads to price drift only for positive sentiment news, and only then if institutional investors are inattentive.

Third, we also interact the type of news (complex versus straightforward), the sentiment of news (positive, negative, or neutral), and the type of investor paying attention (retail versus institutional). We observe that the return drift is stronger for positive financial (straightforward) news when the news is not paid attention to, while the return reversals accompanied with retail attention are stronger for positive business (complex) news. These findings are consistent with the theoretical work of Fedyk (2021), who found that trading volume and price drift are generated in an environment of straightforward news and gradual information diffusion (inattention).

We also extend our analysis to a specific sub-type of retail attention, local retail attention, which is the attention of a subgroup of retail investors who may possess local information advantage and play a different role from generic retail investors. Cziraki, Mondria and Wu (2019) are among the very few papers that have exploited this refinement, finding that local retail attention can be used to forecast future returns, implying that local retail investor's attention is informed attention. Our results are consistent with this notion. We find that local retail attention's role is more similar to institutional attention in that it weakens the return drift and not induces return reversal.

The rest of the paper is organized as follows: Section 3.2 reviews some important studies in the literature; Section 3.3 introduces the dataset and methodology applied in this study; Section 3.4 and 3.5 presents main results and robustness test results. Section 3.6 concludes.

3. 2 Related Literature

Our study relates to two strands of existing literature. First, it relates to studies on the market response to media news. Second, it extends the discussion of the role of different type of investor attention in the context of market response to media news.

Recent studies have documented mixed return response to the sentiment of media news. As expected, most of studies (Tetlock, 2007; Tetlock et al., 2008; Garcia, 2013; Loughran and McDonal, 2011; Engelberg et al., 2012) find that negative return response associates to news with negative sentiment, but the association between positive response and news with positive sentiment lacks statistical significance. Babolmorad and Massoud (2020) argues that the lack of significance on positive news reflects the difficulty of accurately measure positive sentiment as positive words are easily negated in ways that are difficult to classify using traditional bag of word approach. So, Babolmorad and Massoud (2020) proposed a new machine learning approach called Topic-Adaptive Syntax Approach, which measures the sentiment of news considering not only the tone of words, but also word order and contexts. By doing so, they successfully improved the accuracy of the measurement of sentiment and find strong association between positive return response and news with positive sentiment. Their findings highlight the importance of using an accurate approach to identify the sentiment of news.

Our study is mostly similar to Frank and Sanati (2018), which documents the market response to media news and find price overreaction to good news and underreaction to bad news. Different from Frank and Sanati (2018), which define the sentiment of news using the contemporaneous return response, our study contributes to the literature as we apply a textual-analysis approach which could more accurately reflect the sentiment of news and avoid making the assumption that the return response reflects news shock only. In our robustness test, we do find that there is a significant portion of observations that experience inconsistent return response. The inconsistence refers to the case that the sign of return response is inconsistent with the sentiment of news.

Our study also relates to a growing literature discussing the different impacts of investor attention on financial markets from retail and institutional investors. One strand of literature highlights the price-stabilizing role of institutional investors or the price-destabilizing role of retail investors. For example, Barber and Odean (2008) find that attention-grabbing events induce buying pressure from retail investors. Ben-Rephael, Da, and Israelson (2017) show that the post-announcement drifts caused by earning announcement and analyst recommendation change weaken conditional on abnormal attention of institutional, not retail investors. Ben-Rephael, Carlin, Da, and Israelsen (2018) provide evidence that only abnormal institutional attention facilitates price discovery before the filing period of SEC 8-K filings. Chuprinin, Gorbenko, and Kang (2019) document the evidence that abnormal institutional attention improves price correction of mispricing at earning announcement. Da, Hua, Hung, and Peng (2020) find that aggregate firm-level retail attention negatively predicts the market return whereas aggregate institutional attention weakly but positively predicts the market return. They argue that retail attention slows down the incorporation of negative news to the market due to behavioural biases such as disposition effect (Shefrin and Statman, 1985) and ostrich effect (Galai and Sade, 2006).

Another strand of literature concerns the stabilizing role retail attention plays on market efficiency and is most closely related to this study. Liu, Peng, and Tang (2019) find that retail inattention results in lower contemporaneous return response to earning announcement even conditioning on abnormal institutional attention. They use the announcement of important macroeconomic news as an exogenous shock on retail attention to investigate the return response to earning news. Song (2020) looks at the role of retail attention to accounting information during earning announcement period. Consistent with the results in Liu, Peng, and Tang (2019), she finds stronger contemporaneous return reactions and weaker post-announcement drift on earning news when retail investors pay abnormal attention to accounting information. Though the role played by different types of investor attention is inconclusive, we contribute to this literature by introducing the context of market response to media news.

3. 3 Data

The data used in this study comes from multiple sources. The media news are from 15 leading news providers, ranging between January 2014 and December 2018. The news sentiment is identified using Topic-Adaptive Syntax (TASA) Approach proposed in Babolmorad and Massoud (2020). The institutional and retail attention data come from Bloomberg and Google Trends. The stock return and accounting information is from the Center for Research in Securities Prices (CRSP) and Compustat. The sample includes S&P 500 firms as of 2018 that are: i) having at least one media news from at least one of the 15 news providers; ii) not missing market equity or book equity values; iii) with valid institutional and retail attention data; iii) common equities with share code of 10 or 11 in CRSP.

We set the trading date on which the news released as the event date if the release time is during regular trading hours. If the news is released after market close, we assign the next available trading date as the event date. We calculate the abnormal return for stock i on day t around event window n, $AR_{i,t,n}$ making use of the method proposed in Daniel, Grinblatt, Titman, and Wermers (1997) and Wermers (2004).

$$AR_{i,t,n} = R_{i,t,n} - R_{DGTW_i,t,n} \quad (3.1)$$

Where $R_{i,t,n}$ is the stock return for stock i, and $R_{DGTW_i,t,n}$ is the benchmark return of the corresponding $DGTW_i$ group including firms with similar characteristics, on day t, n days away from the media news announcement.

3.3.1 Media News Sentiment

We adopt Topic-Adaptive Syntax (TASA) Approach proposed in Babolmorad and Massoud (2020) to measure the news sentiment. This approach allows us to capture the tone of news headline at granular firm level and also helps us identify the feature of news.

Panel A of Table 3.1 provides details of the identified news sentiment and news feature. In total, there are 130,381 media news items from 15 major news providers in our sample period between 2014 and 2018. Around 45% of news is identified as positive news, but only 15% is identified as negative news. This unbalanced distribution is consistent with the notion that firms have an incentive to manage the sentiment of news through press releases and media tend to report more positive news than negative news on firms which hire investor relations firms (Soloman, 2012; Ahern and Sosyura, 2014). Another advantage of TASA approach is that it also identifies

Table 3. 1: Summary Statistics

This table provides the summary statistics for the data sample between January 2014 and December 2018. Panel A provides the tone of media news identified using the Topic-Adaptive Syntax (TASA) Approach in Babolmorad and Massoud (2020) and the categorization of news. Panel B provides the summary statistics of investor attention and firm characteristics in firm-day level. INST, RETL, and BOTH are indicator variables equals one if there is institutional attention only, retail attention only, and both types of attention for the stock. AR is daily DGTW-adjusted Abnormal Return. ME and BM are market equity and book-to-market ratio. AVOL is daily abnormal trading volume. SDRET is the standard deviation of daily stock returns. HLtH is the ratio of (High - Low)/High of daily price. Spread and Turnover is daily bid-ask spread and daily stock turnover. Panel C provides pairwise Pearson correlation between all the variables.

Panel A: News Tone and News Features Category

	Business	Finance	Business & Finance	Unverified	Total	%Total
Negative	5748	5979	2947	6048	20722	15.89%
Neutral	14097	12282	6154	17564	50097	38.42%
Positive	16532	16735	9453	16842	59562	45.68%
Total	36377	34996	18554	40454	130381	100.00%
%Total	27.90%	26.84%	14.23%	31.03%	100%	

Panel B: Firm Characteristics

Variable	Descriptions	Nobs	Mean	Std	Median	P10	P90
INST	Institutional Attention Only	233,065	17.60%	38.08%	0	0	1
RETL	Retail Attention Only	233,065	4.07%	19.76%	0	0	0
BOTH	Both Institutional and Retail Attention	233,065	2.73%	16.29%	0	0	0
AR	Daily DGTW-adjusted Abnormal Return	233,065	0.009	1.697	-0.006	-1.426	1.456
ME	Market Equity (Million in USD)	233,065	61,235	91,960	24,722	1,441	174,744
BM	Book-to-Market Ratio	233,065	0.560	0.788	0.351	0.106	1.024
AVOL	Abnormal Volume	233,065	1.023	1.296	0.867	0.516	1.598
SDRET	Standard Deviation of Daily Return	233,065	0.017	0.011	0.014	0.008	0.028
HLtH	(High - Low)/High	233,065	0.022	0.016	0.018	0.009	0.040
Spread	Bid-Ask Spread (bps)	233,065	2.5391	4.8636	1.1711	0.4528	5.299
Turnover	Daily Turnover	233,065	0.010	0.020	0.006	0.003	0.020

Panel C: Pearson Correlation between Variables

	AR	Negative	Neutral	Positive	INST	RETL	BOTH	SIZE	BM	AVOL	SDRET	HLtH	Spread
Negative	-0.023												
Neutral	0.003	0.365											
Positive	0.017	0.356	0.449										
INST	0.011	0.110	0.142	0.149									
RETL	0.007	-0.018	-0.019	-0.018	-0.095								
BOTH	0.005	0.094	0.101	0.111	-0.077	-0.035							
ME	-0.001	0.323	0.404	0.390	0.189	-0.039	0.056						
BM	-0.005	-0.048	-0.061	-0.063	0.029	0.002	-0.009	-0.129					
AVOL	-0.068	0.039	0.038	0.046	0.074	0.013	0.162	-0.010	0.003				
SDRET	0.000	-0.050	-0.072	-0.079	-0.076	-0.015	-0.024	-0.217	-0.033	0.083			
HLtH	0.006	-0.016	-0.039	-0.035	0.092	0.016	0.153	-0.199	-0.064	0.281	0.488		
Spread	-0.007	-0.056	-0.084	-0.086	-0.082	0.015	-0.019	-0.213	0.083	0.013	0.297	0.306	
Turnover	0.021	0.020	0.021	0.029	0.074	-0.006	0.169	-0.143	0.004	0.290	0.311	0.439	0.057

the focus of the news story, broken into two broad categories, news focusing on firm fundamentals ("Business") and news focusing on more easily digested information like market price movements ("Finance"). Panel A shows that around 70% of news can be identified as either Business news or Finance news or both. In Section 3.5, we will perform analysis in subsample of different features and explore how investors response differently to different types of news.

3.3.2 Investor Attention Data

We follow Ben-Rephael, Da, and Israelson (2017) and use Bloomberg News Heat – Daily Max Readership to measure institutional attention[17]. The Daily Max Readership data measures the intensity of search and reading activity of Bloomberg users based on the activities in previous 30 days. Ben-Rephael, Da, and Israelson (2017) argue that this measure mostly represents institutional investors' behavior and document evidence that it contains unique information content relevant to institutional investors' behavior and not included in traditional attention measures such as retail attention measures. Specifically, they define institutional attention as an indicator variable equals to one when the intensity of search and reading activity of Bloomberg users is equal to or above the 94th percentile of such activity over the previous 30 days. For more details about this measure, readers are referred to the appendix of Ben-Rephael, Da, and Israelson (2017).

For retail attention, the seminal work by Da, Engelberg, and Gao (2011) shows that Search Volume Index (SVI), which captures the searching intensity of Google users, is a proxy for retail investor attention. Recent studies document refinements by taking advantage of new functions provided in Google Trends. Basistha, Kurov, and Wolfe (2018) find that "related topics" and "related queries" are helpful. Cziraki, Mondria and Wu (2019) find that the regions of search

[17] Ben-Rephael, Da, and Israelson (2017) called this measure as institutional abnormal attention. In our work, we call it institutional attention for brevity purpose as it contains similar meaning.

activity can be used to capture the attention from local retail investors. Zhou (2020) suggests using "investing" subcategory to filter out search activities irrelevant to investor attention (resolve the issue of mistakenly capturing search activities unrelated to investing purpose due to confusing ticker symbols. E.g. excluding searches for cat videos when determining investor attention on the firm Caterpillar, ticker (CAT). We follow Zhou (2020)'s approach and use the "investing" subcategory to identify search activities related to investors attention. We then construct retail attention in a similar fashion as that of institutional attention defining an indicator variable equals one for stocks experiencing higher than 94th percentile search intensity in previous 30 days in Google Trends.

Panel B of Table 3.1 provides summary statistics of attention variable and firm characteristics at Firm-Day level. Note that we categorize attention type into three groups: institutional attention only (INST); retail attention only (RETL); both institutional and retail attention (BOTH). As these are indicator variables, the mean value suggests how frequently a firm receives different type of investor attention. We observe that firms, on average, receive institutional / retail attention without overlapping at 17.6% / 4.1% of total firm-day observations[18]. There are also 2.7% of firm-day observations are identified as receiving both types of investor attention. Other descriptive statistics on firm characteristics show that the firms in our sample are large firms with 61 billion as average market capitalization, whose stock prices are liquid with very small bid-ask spreads.

[18] Note that this is not saying that firms are more likely to receive institutional abnormal attention. The institutional abnormal attention is higher by construction because the variable is calculated in two steps. First, the max readership is constructed in each hour. Second, the daily max readership is the maximum value all the hourly max readership data. So, this way of construction will inflate the frequency of abnormal attention above the nominal 6% level that the 94th percentile ranking would suggest.

Panel C of Table 3.1 shows pairwise Pearson correlation between these variables. Unsurprisingly, positive news is contemporaneously positively correlated with daily abnormal return (0.017) and negative news is contemporaneously negatively correlated with abnormal return (-0.023). Firms' market equity is positively correlated with institutional attention (0.390) but negatively correlated with retail attention (-0.039), suggesting that institutional investors may be more likely to pay attention to larger firms whereas retail investors tend to pay attention to smaller firms.

3. 4 Empirical Results

3.4.1 Univariate Analysis

In this section, we show univariate analysis considering both the news sentiment and the types of investor attention, applying an event-based approach, comparing both contemporaneous return responses and cumulative abnormal returns after the news release. While we find sensible contemporaneous return responses to news, negative returns associated with negative news, positive returns with positive news, we focus our analysis on cumulative abnormal returns after news release to examine the impact of investor attention.

Our primary results are based on a two-stage approach, aggregating the abnormal return from different stocks, and estimating standard errors using the Fama-Macbeth (1973) approach. In the first stage, we calculate the average abnormal return across firms in a bucket n days away from an event day t, $AR_{t,n}$. In the second stage, we compute simple time series average of the abnormal return across all event days, AR_n and obtain heteroskedasticity adjusted standard errors[19].

$$\text{Stage One:} \quad AR_{t,n} = \sum_{i=1}^{I} AR_{i,t,n} \quad (3.2)$$

$$\text{Stage Two:} \quad AR_{n} = \sum_{t=1}^{T} AR_{t,n} \quad (3.3)$$

We find that positive news is followed by a return reversal when only retail investors pay abnormal attention, suggesting a destabilizing role from retail investor attention. In contrast, positive news is not associated with return reversal (or drift) when institutional investors pay abnormal attention. If neither type of investors pays attention, we observe a return drift on positive news. Finally, we don't observe this sharp difference on negative news conditioning on the types of investor attention. Negative news is associated with a larger magnitude contemporaneous return response than we find with positive news when either type of investors pay attention, and no reversal or drift is observed afterwards for negative news.

3.4.1.1 Market responses to different news sentiment and abnormal attention

To compare the impact of news with different sentiment and investor attention, we categorize media news into positive or negative news using the Topic-Adaptive Syntax (TASA) Approach suggested in Babolmorad and Massoud (2020) and we explore how returns vary depending on investor attention and news sentiment. We use four categorizations on attention to news, labelled Noatten, RETL, INST and BOTH. Noatten identifies events that receive neither retail nor institutional attention, RETL identifies events that only receive retail attention, INST identifies events that only receive institutional attention, and BOTH identifies events that receive

[19] This means that abnormal returns AR_{n-1} and AR_n are not independent, but tests of the significance of an individual abnormal return are valid. Our multivariate analysis, presented below, will provide evidence controlling for other conditional effects and joint tests, where appropriate.

both retail and institutional attention. Altogether we then have eight categories of attention and news sentiment, by interacting sentiment (positive and negative) with the four categories of attention.

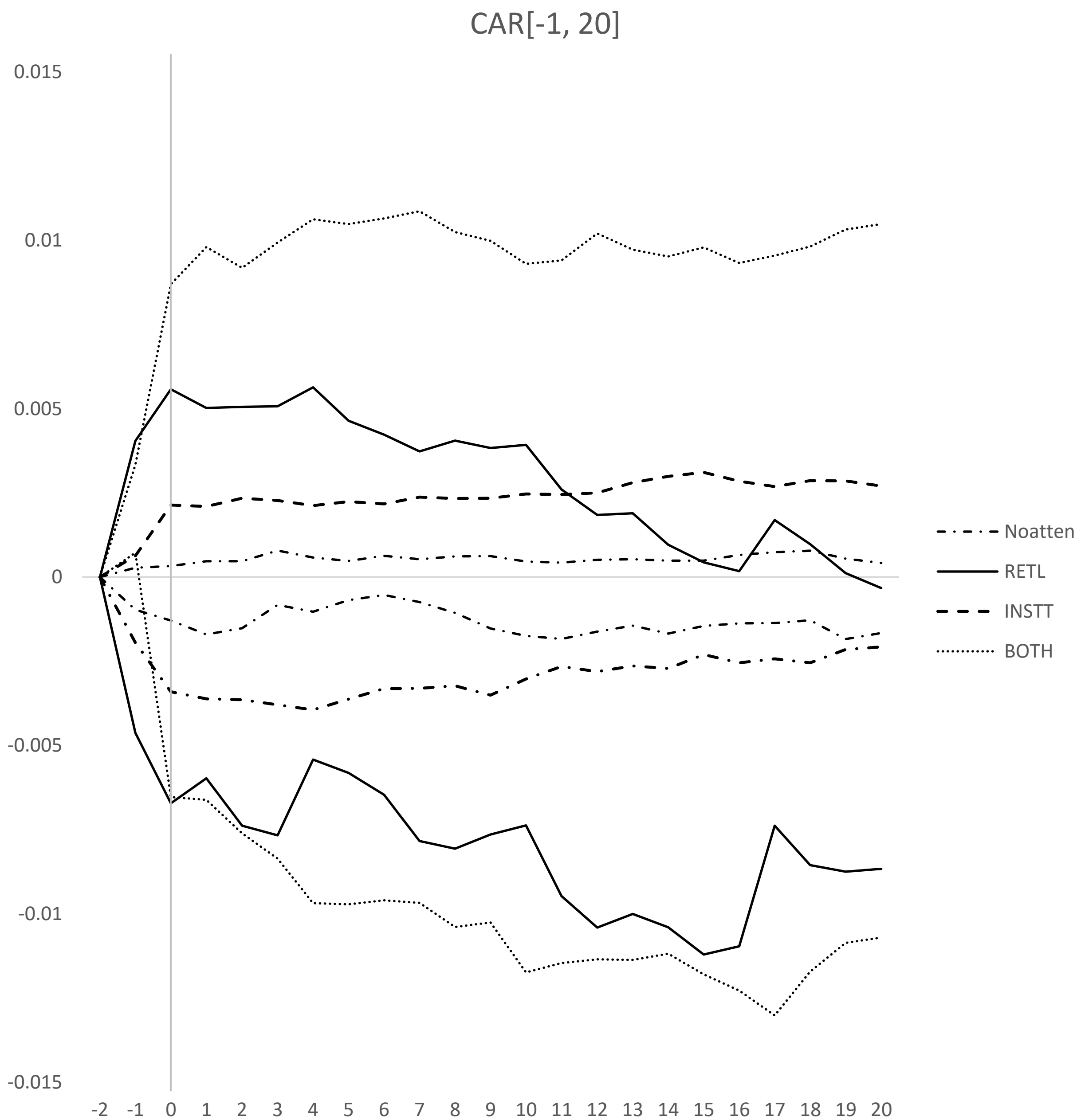
The most interesting findings show up when we look at the trajectories of cumulative abnormal returns. Figure 3.1 illustrates how market responds to media news conditioning on sentiment and attention. We document the CAR starting from one day before to 20 days after news releases. The lines above and below zero are the CAR for positive and negative news respectively. To differentiate the CAR conditioning on the types of attention, we assign the dot dash line, dash line, solid line, and dotted line to the group with no attention (Noatten), institutional attention only (INST), retail attention only (RETL), and both types of attention (BOTH) [20].

The most notable findings are that retail attention induce price overreaction on positive news, suggesting a destabilizing role of retail investors. The solid line in the upper half of Figure 1, depicting the CAR for positive news only receives high retail attention, exhibits a reverse U-shape. When news is released, it first draws strong contemporaneous return response, then the CAR will reverse gradually to a level as low as that of news without attention. Untabulated results show that the CAR [1,20] is -59 basis points (t-stats=2.002), which is both statistically and economically significant. This return reversal documents a destabilizing role of retail attention on media news. It is consistent with the findings in Frank and Sanati (2018) that positive return shock is followed by return reversals with consideration of retail investors attention habits.

[20] The CAR is calculated based on daily DGTW-adjusted abnormal returns. To highlight the comparison between groups, the CAR is provided in a relative sense by subtracting the CAR of a benchmark group without media news and abnormal attention.

Figure 3. 1: Cumulative Abnormal Return and News Sentiment and Abnormal Attention

Figure 3.1 illustrates the market responses to news with different sentiment and with different types of abnormal attention. This figure draws the trajectory of DGTW-adjusted CAR in the window of [-1, 20]. The dot-dash line, dash line, solid line, and dotted line represent the CAR [-1,20]] with no attention (Noatten), institutional attention only (INST), retail attention only (RETL), and both attention (BOTH). The lines in the upper and lower half represent the CAR [-1,20] for the media news identified as positive and negative respectively. The abnormal return in CAR [-1,20] is calculated in a two-stage approach described in section 4.1.1.



In addition, we find that positive information is quickly incorporated into prices upon news announcement once institutional investors pay attention, suggesting a stabilizing role of institutional investors. The dash line and dotted line in the upper half, which represent the condition that only institutional investors or both types of investors pay abnormal attention, experience a sudden jump during [-1, 0] window and maintains at a high level without much fluctuation afterward. In other words, there are large contemporaneous market responses, and no further return drifts conditioning on institutional abnormal attention. Untabulated results show that the CAR [1,20] is only 6 basis points (t-stats =0.658) and 18 basis points (t-stats = 0.532) respectively without statistical significance.

Similar to the findings on positive news, we observe strong contemporaneous responses and no further drifts to negative news conditioning on institutional attention. However, in contrast to positive news, we don't observe price overreaction when only retail investors pay attention to negative news. This is consistent with the notion of asymmetric patterns from the studies about retail investors trading behaviors. That is, retail attention is more likely to induce buying pressure than selling pressure because they are less likely involved in short selling behaviors. (Baber and Odean, 2008; Cziraki, Mondria, and Wu, 2019)

3.4.1.2 Further Analysis

Table 3.2 provides a closer look of the return responses in the window of [-3, 3] around media news. Panel A presents the abnormal return to news with positive sentiment. The results in bucket 0 show that the contemporaneous market response conditions on the type of investor attention. When the news doesn't receive attention from either retail or institutional investors, there is insignificant abnormal return during the day news announced. But the return responses are much stronger if at least one type of investors pay abnormal attention. For example, news receiving retail

Table 3. 2: DGTW-Adjusted Abnormal Return

This table provides the market responses to media news. The DGTW-Adjusted Abnormal Return is calculated using two-stage approach discussed in section 4.1.1. AR[-3], ..AR[3] are abnormal returns calculated as AR_n in equation (3.3). The t-stats are Newey-West standard errors that is adjusted for heteroskedasticity and autocorrelation correction. Panel A and Panel B shows the results on news with positive and negative tone respectively. Noatten, RETL, INST, and BOTH presents groups of news with no attention, retail attention only, institutional attention only and both types of attention.

Panel A: DGTW-Adjusted Abnormal Return for Positive News

Sentiment	Attention	Trading Days	DGTW-Adjusted Abnormal Return						
			AR [-3]	AR [-2]	AR [-1]	AR [0]	AR [1]	AR [2]	AR [3]
One or More Positive News	Noatten	1223	0.02%	0.05%***	0.01%	0.00%	0.02%*	0.01%	0.03%***
			(1.48)	(3.09)	(1.22)	(0.00)	(1.73)	(0.81)	(3.06)
	RETL	679	0.13%***	0.15%*	0.41%***	0.16%**	-0.02%	0.11%*	0.06%
			(2.90)	(1.72)	(2.67)	(2.01)	(-0.28)	(1.70)	(1.30)
	INST	1210	0.05%***	0.01%	0.07%***	0.15%***	0.00%	0.03%*	0.00%
			(2.70)	(0.54)	(3.14)	(6.28)	(-0.02)	(1.90)	(0.01)
	BOTH	898	0.14%***	0.12%*	0.35%***	0.57%***	0.06%	-0.05%	0.04%
			(2.76)	(1.83)	(3.33)	(4.24)	(0.83)	(-0.77)	(0.74)

Panel B: DGTW-Adjusted Abnormal Return for Negative News

Sentiment	Attention	Trading Days	DGTW-Adjusted Abnormal Return						
			AR [-3]	AR [-2]	AR [-1]	AR [0]	AR [1]	AR [2]	AR [3]
One or More Negative News	Noatten	1210	-0.03%	-0.04%*	-0.10%***	-0.04%***	0.00%	0.02%	0.07%***
			(-1.59)	(-1.79)	(-4.72)	(-2.63)	(-0.26)	(1.07)	(3.90)
	RETL	309	0.14%*	-0.13%	-0.35%**	-0.19%	-0.02%	-0.14%	0.02%
			(1.75)	(-1.20)	(-1.96)	(-1.38)	(-0.19)	(-1.42)	(0.32)
	INST	1165	-0.02%	-0.05%**	-0.15%***	-0.15%***	-0.01%	0.01%	-0.01%
			(-0.80)	(-2.16)	(-5.14)	(-4.76)	(-0.49)	(0.41)	(-0.52)
	BOTH	669	0.10%**	0.15%***	-0.04%	-0.69%***	-0.02%	-0.03%	-0.05%
			(1.97)	(2.82)	(-0.37)	(-4.35)	(-0.28)	(-0.47)	(-0.94)

or institutional abnormal attention are associated with 16 or 15 basis points abnormal returns. If the news receives both types of attention, the return response is over three times as large, 57 basis points. Another finding is that investor attention also impacts the return response after a news announcement. There is a strong statistically significant drift when there is no attention to positive news, though there is some evidence of drift even with attention. We will provide additional evidence below to explore if these patterns are economically significant and robust to controls.

Panel B of Table 3.2 presents the daily abnormal return to news with negative sentiment. We find that the impact of attention on return response is similar to positive news. When news receives neither retail nor institutional attention, the contemporaneous return response is the weakest in magnitude. The return response is larger if at least one type of investors pays attention. Not surprisingly, when both types of investors pay abnormal attention to news, there is as large as 70 basis point daily contemporaneous abnormal return. There is some evidence of reversal for the case of investor inattention, but as we will explore, this largely disappears when we introduce controls.

3.4.2 Multivariate Analysis

In this section, we perform panel regression analysis to tease out marginal effects of news sentiment on financial markets conditioning on different types of attention while controlling for a broad range of control variables. We first explore a very high-level view of the response of prices to news and attention then break down these price responses to explore the impact of different types of investors, different types of news (both valence and focus), and different informativeness of investor attention. This analysis will help us understand conflicting results in the literature on the impact of investor attention on market prices.

3.4.2.1 Pure vs Mixed sentiment

To start with, we perform an analysis to explore the impact on market prices of pure news sentiment and overall investor attention. We estimate the following regression,

$$CAR_{i,t} = \alpha + \beta_1 * Pure_{i,t} + \beta_2 * Pure_{i,t} * Atten_{\mathrm{i,t}} + \gamma_1 * Mix_{i,t}$$

$$+\gamma_2 * Mix_{i,t} * Atten_{\mathrm{i,t}} + \delta_1 * Atten_{\mathrm{i,t}} + \vartheta_1 * Controls_{i,t} + \varepsilon_{i,t} \quad (3.4)$$

where $CAR_{i,t}$ is cumulative abnormal return for firm i at date t. $Pure_{i,t}$ is a sentiment index variable which equals one (negative one) when there is only positive (negative) news for firm i at date t, and zero otherwise[21]. $Mix_{i,t}$ is an indicator variable equals one if there are mixed tone of news, namely both positive and negative news, for firm i at date t and zero otherwise. $Atten_{\mathrm{i,t}}$ is an indicator variable equals one if there is institutional and/or retail abnormal attention for firm i at date t and zero otherwise. $Controls_{i,t}$ contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). Table 3.3 provides the estimation results for equation (4) based on full sample period between Jan 1st, 2014 and Dec 31st, 2018. The analysis is performed in the windows of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20] around the event dates.

We find evidence that investor attention facilitates information incorporation of media news into asset prices. In the windows of [1, 5], [1, 10], [1, 15], and [1, 20] we observe that a pure sentiment to news, when associated with investor inattention, is followed by a statistically

[21] We ignore the neutral news by assuming that neutral news have insignificant effects on stock returns. This assumption is supported in the untabulated results that there is no significant effects of news with neutral tone.

Table 3. 3: The Effect of Pure vs Mix Sentiment News Conditioning on Abnormal Attention

This table provides estimation results of panel regression analysis for equation (3.4) in the sample period between Jan 1st, 2014 and Dec 31st, 2018. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Pure is a sentiment index equals one and negative one if there is only positive or negative news and zero otherwise. Mix is an indicator variable equals one if there are both positive and negative news announced for the same firm in the same date. Atten is an indicator variable equals on if there is either institutional or retail attention for the related firm in the same date as the news released. The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1) **[-1, 0]**	**(2)** **[1, 5]**	**(3)** **[1, 10]**	**(4)** **[1, 15]**	**(5)** **[1, 20]**
Pure	0.059***	0.046*	0.078*	0.142**	0.166**
	(3.81)	(1.66)	(1.81)	(2.48)	(2.49)
Pure*Atten	0.460***	0.037	-0.009	-0.075	-0.093
	(8.21)	(1.00)	(-0.16)	(-1.00)	(-1.23)
Mix	-0.026	0.037	0.010	-0.036	-0.011
	(-0.81)	(0.43)	(0.06)	(-0.17)	(-0.04)
Mix*Atten	-0.222***	-0.027	-0.022	0.072	0.041
	(-2.78)	(-0.32)	(-0.19)	(0.46)	(0.22)
Atten	0.063***	-0.003	0.042	0.042	0.051
	(2.80)	(-0.14)	(1.07)	(0.80)	(0.92)
Ln(ME)	-0.002	0.005	0.010	0.018	0.030
	(-0.33)	(0.35)	(0.38)	(0.46)	(0.59)
BM	-0.007	-0.063***	-0.123***	-0.187***	-0.243***
	(-1.13)	(-4.98)	(-5.14)	(-5.35)	(-5.28)
AVOL	-0.129**	-0.010	-0.020	-0.002	-0.003
	(-2.12)	(-1.07)	(-1.02)	(-0.07)	(-0.11)
SDRET	-1.385	-0.219	-1.722	3.132	11.260
	(-0.93)	(-0.08)	(-0.30)	(0.36)	(1.02)
HLtH	4.506*	-0.478	0.553	-4.139	-6.496
	(1.73)	(-0.29)	(0.25)	(-1.29)	(-1.59)
Spread	0.0017	0.0161**	0.0301*	0.0504**	0.0657**
	(0.64)	(2.15)	(1.90)	(2.16)	(2.20)
Turnover	-1.453	0.473	2.305	3.385	1.414
	(-1.44)	(0.18)	(0.29)	(0.27)	(0.09)
CAR[-5,-2]	17.153***	-1.422***	-1.844**	-1.375	-1.743
	(54.56)	(-2.75)	(-2.43)	(-1.48)	(-1.37)
Constant	0.037	0.015	-0.003	-0.059	-0.210
	(0.32)	(0.09)	(-0.01)	(-0.12)	(-0.32)
Observations	219,487	219,487	219,487	219,487	219,487
Adj R^2	0.091	0.002	0.003	0.003	0.004
DATE FE	Yes	Yes	Yes	Yes	Yes

significant return persistence, suggesting a delay of information transmission. Specifically, the CAR has a gradual increase from 4.6 basis points to as high as 16.6 basis points from [1, 5] to [1, 20] window. However, the CAR after a pure sentiment news release is statistically insignificant when there is investor attention. This is consistent with the idea that investor attention facilitates information incorporation and removes return persistence. This finding is analogous to that of Ben-Rephael, Da, and Israelson (2017), which documents that the post earning announcement drifts weaken conditioning on institutional attention.

Also, we find that both the sentiment measure and the investor attention measure work sensibly on contemporaneous price movements. The results on CAR [-1, 0] show that a pure sentiment news is associated with a consistent contemporaneous return response, and the magnitude of the response also conditions on whether there is investor attention. Specifically, when there is no attention, a pure sentiment news is associated with a 5.9 basis points daily abnormal return within the window of [-1, 0]. The positive sign means that positive news receives positive responses and negative news receive negative responses[22], suggesting that the sentiment of news is incorporated into asset prices, consistent with the findings in Babolmorad and Massoud (2020). In addition, there is an additional 46 basis points return response if there is investor attention on the news day.

It is also interesting to investigate the effects of mixed sentiment news. We observe that there is a small 2.6 basis points negative return response if there are both positive and negative news in the same day, suggesting that the effects of negative and positive sentiment would cancel each other on average. Interestingly, if there is investor attention during the day of news release,

[22] The attention and price response are contemporaneous and hence we can make no firm conclusions that one drives the other. For this reason, most of our analysis focuses on windows of time following the news event.

we find an additional negative 22 basis points of abnormal return, implying that negative news dominants positive news when it occurs with investor attention. This conflicting news is incorporated into prices quickly, as there is no statistically or economically significant return drift after the news event day.

These results give us our first indications of the importance of both attention from investors and unambiguous news sentiment, and they help us begin to understand the delicacy of price responses; news is only impactful if it is paid attention to, and if there are conflicting news, displaying on the same day both positive and negative sentiment, the impact on prices is similar to what it would be with only negative news valence.

3.4.2.2 Institutional vs Retail Attention

There is a growing literature exploring the role of retail investor attention on financial markets. Hence we are interested in separating out the effects of attention from different types of investors. Is attention from retail investors similarly impactful as institutional investors? This can be investigated by simply splitting the attention (Atten) indicator variable into institution only (INST), retail only (RETL), and both attention (BOTH), and updating the specification as below:

$$CAR_{i,t} = \alpha + \beta_1 * Pure_{i,t} + \beta_2 * Pure_{i,t} * INST_{\mathrm{i,t}} + \beta_3 * Pure_{i,t} * RETL_{\mathrm{i,t}} + \beta_4 * Pure_{i,t} * BOTH_{\mathrm{i,t}} + \gamma_1 * Mix_{i,t} + \gamma_2 * Mix_{i,t} * INST_{\mathrm{i,t}} + \gamma_3 * Mix_{i,t} * RETL_{\mathrm{i,t}} + \gamma_4 * Mix_{i,t} * BOTH_{\mathrm{i,t}} + \delta_1 * INST_{\mathrm{i,t}} + \delta_2 * RETL_{\mathrm{i,t}} + \delta_3 * BOTH_{\mathrm{i,t}} + \vartheta_1 * Controls_{i,t} + \varepsilon_{i,t} \tag{3.5}$$

where $INST_{\mathrm{i,t}}$ is an indicator variable equaling one if there is only institutional attention on firm i at date t and zero otherwise, $RETL_{\mathrm{i,t}}$ is an indicator variable equaling one if there is only retail attention on firm i at date t and zero otherwise and $BOTH_{\mathrm{i,t}}$ is an indicator variable equaling one if there are both institutional and retail attention on firm i at date t and zero otherwise. Table 3.4

provides the estimation on equation (3.5) based on full sample period between Jan 1st, 2014 and Dec 31st, 2018.

Consider first retail attention. In the case that there is only retail attention, we observe an additional 48 basis points during [-1, 0] window on pure sentiment index. Interestingly, we find a conditional reversal of 49 basis points during [1, 20] window, suggesting that the retail attention will first induce a price overreaction with the news release, then the price will fully reverse. This is a striking result showing that retail investors' attention may play a destabilizing role on financial markets. It is consistent with the studies documenting retail attention's destabilizing role (Barber and Odean, 2008; Da, Engelberg, and Gao, 2011), but inconsistent with those documenting stabilizing role for retail attention (Liu, Peng, and Tang, 2019; Zhou, 2020). We will explore further analysis to show much of this reversal is due to over-reaction to positive news.

In contrast to retail attention, the institutional attention has associated with much smaller reversals (less than 10 bps), suggesting that institutional attention stabilizes financial markets. These results also suggest that the effect of aggregated attention reported in Table 3.3 are driven by institutional investors, and should we neglect to split investor classes into institutional and retail we are likely to obscure important price reactions to news and investor attention.

3.4.2.3 Positive vs Negative Sentiment

The above analysis helps us stylize important effects from investor attention to news but obscures potential variation between positive and negative news sentiment. Recent studies show that there are asymmetric impacts between positive and negative shocks on financial markets (Frank and Sanati, 2008). In addition, studies on retail attention also document that there are

Table 3. 4: The Effect of Pure vs Mix Sentiment News Conditioning on Attention Types

This table provides estimation results of panel regression analysis for equation (3.5) in the sample period between Jan 1st, 2014 and Dec 31st, 2018. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Pure is a sentiment index equals one and negative one if there is only positive or negative news and zero otherwise. Mix is an indicator variable equals one if there are both positive and negative news announced for the same firm in the same date. INST is an indicator variable equals one if there is only institutional abnormal attention and zero otherwise. RETL is an indicator variable equals one if there is only retail abnormal attention and zero otherwise. BOTH is an indicator variable equals one if there are both institutional and retail abnormal attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1) [-1, 0]	**(2) [1, 5]**	**(3) [1, 10]**	**(4) [1, 15]**	**(5) [1, 20]**
Pure	0.058***	0.045	0.077*	0.141**	0.165**
	(3.73)	(1.62)	(1.78)	(2.46)	(2.47)
Pure*INST	0.270***	0.013	-0.027	-0.081	-0.095
	(6.43)	(0.32)	(-0.46)	(-1.03)	(-1.15)
Pure*RETL	0.480***	0.013	-0.144	-0.377*	-0.491*
	(3.17)	(0.11)	(-0.88)	(-1.83)	(-1.96)
Pure*BOTH	1.448***	0.216	0.222	0.207	0.223
	(8.02)	(1.52)	(1.36)	(1.07)	(1.26)
Mix	-0.028	0.035	0.007	-0.038	-0.014
	(-0.87)	(0.40)	(0.05)	(-0.18)	(-0.05)
Mix*INST	-0.156**	-0.007	0.016	0.119	0.065
	(-2.41)	(-0.09)	(0.14)	(0.73)	(0.32)
Mix*RETL	-0.514**	-0.120	-0.043	-0.471	-0.453
	(-2.02)	(-0.40)	(-0.13)	(-1.12)	(-0.90)
Mix*BOTH	-0.301	0.001	-0.035	0.203	0.227
	(-1.48)	(0.00)	(-0.15)	(0.71)	(0.70)
INST	0.060**	-0.011	0.038	0.053	0.048
	(2.19)	(-0.36)	(0.78)	(0.83)	(0.68)
RETL	0.118***	0.050	0.111	0.103	0.138
	(3.05)	(1.08)	(1.48)	(1.24)	(1.39)
BOTH	-0.018	-0.074	-0.091	-0.198	-0.125
	(-0.18)	(-0.80)	(-0.72)	(-1.50)	(-0.83)
Observations	219,487	219,487	219,487	219,487	219,487
Adj R^2	0.092	0.002	0.003	0.003	0.004
Controls	Yes	Yes	Yes	Yes	Yes
DATE FE	Yes	Yes	Yes	Yes	Yes

asymmetric impacts of retail attention (Barber and Odean, 2008; Hartzmark, 2014) on positive and negative news.

In this section, we look at the effects of positive and negative news separately. We look at the effect of pure positive news that is not overlapped with negative news announcement. For brevity, we look at negative news that pools pure negative news and mixed sentiment together given that negative news dominants positive news, as shown in previous results[23]. The specification is,

$$CAR_{i,t} = \alpha + \beta_1 * Negative_{i,t} + \beta_2 * Negative_{i,t} * INST_{\mathrm{i,t}} + \beta_3 * Negative_{i,t} * RETL_{\mathrm{i,t}} + \beta_4 * Negative_{i,t} * BOTH_{\mathrm{i,t}} + \gamma_1 * Pure\ Positive_{i,t} + \gamma_2 * Pure\ Positive_{i,t} * INST_{\mathrm{i,t}} + \gamma_3 * Pure\ Positive_{i,t} * RETL_{\mathrm{i,t}} + \gamma_4 * Pure\ Positive_{i,t} * BOTH_{\mathrm{i,t}} + \delta_1 * INST_{\mathrm{i,t}} + \delta_2 * RETL_{\mathrm{i,t}} + \delta_3 * BOTH_{\mathrm{i,t}} + \vartheta_1 * Controls_{i,t} + \varepsilon_{i,t} \quad (3.6)$$

where $Negative_{i,t}$ is an indicator variable equals one if there is negative news announced for firm i at date t and zero otherwise. $Pure\ Positive_{i,t}$ is an indicator variable equals one if there is positive news and no negative news announced for firm i at date t and zero otherwise. Table 3.5 provides the estimation on equation (3.6) based on full sample period between Jan 1st, 2014 and Dec 31st, 2018.

Consistent with previous studies, we observe asymmetric return responses to positive and negative news. We find that the previously documented return drifts and reversals are mainly driven by pure positive news. There is a 5 basis points contemporaneous return response for positive news in window [-1, 0], and a 21 basis points return drift in window [1, 20] without

---

[23] Untabulated results show that the effects of the pure negative news are similar to the pooled negative news. We pool pure negative news and mixed news only for brevity.

Table 3. 5: The Effect of Negative vs Pure Positive Sentiment Conditioning on Attention Types

This table provides estimation results of panel regression analysis for equation (3.6) in the sample period between Jan 1st, 2014 and Dec 31st, 2018. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Negative is an indicator variable equals one if there is negative news released and zero otherwise. Pure Positive is an indicator variable equals one if there is only positive news announced for the same firm in the same date. INST is an indicator variable equals one if there is only institutional attention and zero otherwise. RETL is an indicator variable equals one if there is only retail attention and zero otherwise. BOTH is an indicator variable equals one if there are both institutional and retail attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	[-1,0]	**[1, 5]**	**[1, 10]**	**[1, 15]**	**[1, 20]**
Negative	-0.113***	0.055	0.003	-0.015	0.007
	(-3.13)	(1.02)	(0.03)	(-0.12)	(0.04)
Negative*INST	-0.444***	-0.133	-0.024	0.010	-0.126
	(-4.98)	(-1.26)	(-0.19)	(0.06)	(-0.64)
Negative*RETL	-0.806**	-0.215	-0.275	-0.489	-0.433
	(-2.57)	(-0.65)	(-0.67)	(-0.94)	(-0.66)
Negative*BOTH	-1.931***	-0.064	-0.035	0.146	-0.069
	(-6.66)	(-0.24)	(-0.11)	(0.42)	(-0.22)
Pure Positive	0.050***	0.068*	0.100*	0.178**	0.210**
	(2.94)	(1.95)	(1.73)	(2.27)	(2.29)
Pure Positive*INST	0.250***	-0.021	-0.041	-0.095	-0.155
	(5.48)	(-0.37)	(-0.54)	(-1.04)	(-1.53)
Pure Positive*RETL	0.479***	0.006	-0.182	-0.455**	-0.552**
	(3.15)	(0.05)	(-1.06)	(-2.03)	(-2.01)
Pure Positive*BOTH	1.362***	0.216	0.233	0.227	0.186
	(6.21)	(1.35)	(1.15)	(0.93)	(0.77)
INST	0.061**	-0.001	0.042	0.063	0.068
	(2.23)	(-0.04)	(0.82)	(0.96)	(0.93)
RETL	0.113***	0.051	0.120	0.113	0.148
	(2.87)	(1.04)	(1.56)	(1.34)	(1.47)
BOTH	0.017	-0.081	-0.111	-0.201	-0.097
	(0.15)	(-0.83)	(-0.82)	(-1.37)	(-0.62)
Observations	219,487	219,487	219,487	219,487	219,487
Adjusted R-squared	0.093	0.002	0.003	0.003	0.004
Controls	Yes	Yes	Yes	Yes	Yes
DATE FE	Yes	Yes	Yes	Yes	Yes

attention, both strongly statistically significant. Conditioning on retail attention, we find an additional 48 basis points return response when the news released, which is followed by a reversal of over 55 basis points. These results suggests that retail attention destabilizes prices by inducing a price overreaction, similar to what we observe looking at pure sentiment, but now identifying this as associated with only positive sentiment.

3.4.2.4 Subsample Analysis for Different News Features: Business vs Finance News

We further investigate the effects of sentiment and investor attention from different features of news, with a categorization meant to identify differing market responses to news that is complex and news that is relatively simple to digest. The Topic-Adaptive Syntax (TASA) Approach enables us to categorize news into business related (complex) news and finance related (simple) news. According to Babolmorad and Massoud (2020), business news are those discussing about firm's operations and management that are beyond the analysis of stock prices, such as launching new products. While finance news are those directly reporting the trading or fundamental financial information of stocks, such as discussion of earning reports. Untabulated results show that 14.2% (12.5%) of positive (negative) finance news are being paid retail abnormal attention, which is more frequently than that received by positive and negative business news, at 9.1% and 10.2% each. In addition, untabulated results show that negative news receives more attention than positive news, regardless of the type of investors and the features of news. Table 3.6 provides the estimation of the extended analysis on Negative and Pure Positive News by splitting news by features.

First, we observe statistically and economically significant return reversals only for the case of positive business news when accompanied with retail abnormal attention. Second, we find

Table 3. 6: The Effect of Business vs Finance Sentiment Conditioning on Attention Types

This table provides the estimation of the extended analysis on Negative and Pure Positive News by splitting news by features in the sample period between Jan 1st, 2014 and Dec 31st, 2018. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Negative_Bus / Negative_Fin is an indicator variable equals one if there is negative business/finance news released and zero otherwise. Pure Positive_Bus / Pure Positive_Fin is an indicator variable equals one if there is only positive business / finance news announced for the same firm in the same date. INST is an indicator variable equals one if there is only institutional attention and zero otherwise. RETL is an indicator variable equals one if there is only retail attention and zero otherwise. BOTH is an indicator variable equals one if there are both institutional and retail attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1)	(2)	(3)	(4)	(5)
	[-1,0]	[1, 5]	[1, 10]	[1, 15]	[1, 20]
Negative_Bus	-0.016	-0.079	-0.056	0.012	0.047
	(-0.51)	(-1.20)	(-0.50)	(0.08)	(0.23)
Negative_Bus*INST	-0.182***	0.033	0.137	0.131	0.032
	(-2.63)	(0.36)	(1.00)	(0.80)	(0.14)
Negative_Bus*RETL	-0.775**	-0.511*	-0.482	-0.387	-0.404
	(-2.52)	(-1.68)	(-1.28)	(-0.72)	(-0.61)
Negative_Bus*BOTH	-0.394	0.125	-0.292	-0.043	0.028
	(-1.26)	(0.56)	(-0.96)	(-0.13)	(0.07)
Negative_Fin	-0.071*	0.169**	-0.010	-0.044	-0.031
	(-1.72)	(2.08)	(-0.07)	(-0.22)	(-0.13)
Negative_Fin*INST	-0.308***	-0.087	0.034	0.047	0.090
	(-3.24)	(-0.87)	(0.25)	(0.25)	(0.41)
Negative_Fin*RETL	-0.301	0.094	-0.288	-0.674	-0.475
	(-0.93)	(0.14)	(-0.60)	(-1.57)	(-0.80)
Negative_Fin*BOTH	-0.643***	-0.411	-0.397	-0.229	-0.432
	(-2.64)	(-1.61)	(-1.42)	(-0.76)	(-1.17)
Pure Positive_Bus	-0.013	0.040	0.111	0.180*	0.161
	(-0.51)	(0.87)	(1.54)	(1.66)	(1.30)
Pure Positive_Bus *INST	0.065	-0.093	-0.021	-0.016	-0.032
	(1.00)	(-1.16)	(-0.18)	(-0.11)	(-0.20)
Pure Positive_Bus *RETL	0.273	-0.156	-0.310	-0.856**	-0.994**
	(1.22)	(-0.59)	(-0.92)	(-1.99)	(-2.24)
Pure Positive_Bus *BOTH	0.455	0.355	0.002	-0.107	0.002
	(1.46)	(1.46)	(0.01)	(-0.24)	(0.00)
Pure Positive_Fin	0.060**	0.132***	0.192***	0.207**	0.240**
	(2.00)	(2.74)	(2.88)	(2.22)	(2.16)
Pure Positive_Fin*INST	0.194***	0.029	-0.125	-0.154	-0.154

	(3.37)	(0.32)	(-0.96)	(-0.97)	(-1.00)
Pure Positive_Fin *RETL	0.249	-0.385*	-0.568*	-0.563	-0.585
	(1.10)	(-1.87)	(-1.76)	(-1.54)	(-1.35)
Pure Positive_Fin *BOTH	1.453***	-0.034	0.007	0.075	-0.074
	(4.56)	(-0.14)	(0.03)	(0.24)	(-0.21)
INST	0.079***	-0.003	0.038	0.050	0.050
	(2.83)	(-0.11)	(0.73)	(0.76)	(0.67)
RETL	0.123***	0.060	0.128*	0.124	0.152
	(2.97)	(1.24)	(1.67)	(1.44)	(1.48)
BOTH	0.018	-0.054	-0.038	-0.131	-0.040
	(0.17)	(-0.55)	(-0.28)	(-0.94)	(-0.25)
Observations	219,487	219,487	219,487	219,487	219,487
Adjusted R-squared	0.094	0.002	0.003	0.003	0.004
Controls	Yes	Yes	Yes	Yes	Yes
DATE FE	Yes	Yes	Yes	Yes	Yes

statistically and economically significant return drift only for the case of positive finance news when accompanied by a lack of investor (retail and institutional) attention.

These results are consistent with the findings in Fedyk (2021), who documents that gradual information diffusion and investor inattention drives the trading for straightforward (finance related) news. In our case, the gradual information diffusion, due to lack of attention, results in return persistence on positive finance news. These results also present a new stylized fact, that retail attention in absence of institutional attention is destabilizing only if the news is complex, and highlights the importance of controlling not only for who is paying attention, but who else is paying attention and how complex the information environment is. Analysis of market reaction that does not control for these covariates can easily flip results and obscure the relationship between attention, news and market responses.

3.4.2.5 Subsample Analysis for Local Retain Attention

Strands of the literature have documented a potentially stabilizing role for retail attention, inconsistent with much of the rest of the literature and inconsistent with the evidence we have provided so far here. In this section, we focus on the analysis of the impact of attention by including another type of retail investors, local retail investors, which is defined as the attention paid by retail investors who is in the same state as that of the firms headquarter. Previous studies (Cziraki et al., 2019; Zhou, 2020) show that local attention could be informed attention. In order to pull apart potentially conflicting impacts from informed and uniformed retail investors, we extend our analysis to substitute local retail investor attention with that of local versus national retail investor attention. Table 3.7 provides the estimation of the extended analysis by including these types of investors. We collapse the category of joint attention from retail and institutional investors into the institutional investor category, as previous results indicate that institutional investor attention

Table 3. 7: The Effect of Sentiment Conditioning on Attention (With Local Retail Attention)

This table provides estimation results of panel regression analysis in the sample period between Jan 1st, 2014 and Dec 31st, 2018. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Negative is an indicator variable equals one if there is negative news released and zero otherwise. Pure Positive is an indicator variable equals one if there is only positive news announced for the same firm in the same date. Institutional is an indicator variable equals one if there is institutional attention and zero otherwise. Local Retail is an indicator variable equals one if there is local retail attention and zero otherwise. National Retail is an indicator variable equals one if there is national retail attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1)	(2)	(3)	(4)	(5)
	[-1,0]	[1, 5]	[1, 10]	[1, 15]	[1, 20]
Negative	-0.108***	0.053	-0.015	-0.022	-0.005
	(-2.74)	(0.96)	(-0.14)	(-0.16)	(-0.03)
Negative*Institutional	-0.721***	-0.158	-0.038	0.020	-0.132
	(-7.11)	(-1.39)	(-0.27)	(0.11)	(-0.67)
Negative*Local Retail	-0.005	0.269	-0.100	-0.360	-0.369
	(-0.03)	(0.74)	(-0.26)	(-0.75)	(-0.70)
Negative*National Retail	-0.762**	-0.174	0.005	-0.073	0.227
	(-2.21)	(-0.44)	(0.01)	(-0.12)	(0.30)
Pure Positive	0.044**	0.068*	0.108*	0.185**	0.214**
	(2.32)	(1.79)	(1.71)	(2.17)	(2.10)
Pure Positive*Institutional	0.410***	0.007	-0.010	-0.050	-0.098
	(6.39)	(0.12)	(-0.12)	(-0.47)	(-0.85)
Pure Positive*Local Retail	0.185	-0.094	-0.199	-0.134	0.013
	(1.34)	(-0.81)	(-1.06)	(-0.59)	(0.05)
Pure Positive*National Retail	0.414***	-0.087	-0.255	-0.612**	-0.650*
	(3.09)	(-0.57)	(-1.33)	(-2.20)	(-1.91)
Institutional	0.056	-0.008	0.013	0.026	0.051
	(1.53)	(-0.26)	(0.23)	(0.36)	(0.61)
Local Retail	0.113**	0.048	0.044	0.071	0.017
	(2.43)	(0.96)	(0.60)	(0.85)	(0.17)
National Retail	0.065*	0.037	0.130	0.100	0.122
	(1.94)	(0.70)	(1.58)	(1.07)	(1.07)
Observations	181,314	181,314	181,314	181,314	181,314
Adjusted R-squared	0.094	0.002	0.002	0.002	0.003
Controls	Yes	Yes	Yes	Yes	Yes
DATE FE	Yes	Yes	Yes	Yes	Yes

overwhelms any impact from retail attention on market prices. Untabulated results indicate that our main results are insensitive to this collapsed specification.

We find that the impact of local retail attention is similar to institutional attention, but in sharp contrast to the national retail attention. That is, local retail attention is not associated with a return reversal while in contrast, national retail abnormal attention results in a large economic and statistically significant reversal of over 60 basis points of CAR on positive news.

3. 5 Robustness

3.5.1 Sentiment Measure: Textual-Based vs Return Shock

While we use a direct measure of the sentiment of the news, several papers have proxied for this sentiment by using the sign of the return when the news is released (see, for instance, Frank and Sanati, 2018; Ma, Xiong, and Feng, 2020). In this section, we explore how our textual-based approach compares to this return shock approach.

Table 3.8 provides the estimation results using the specification outlined in equation (3.6) by measuring the sentiment using return shock approach. Negative (Positive) is an indicator variable equals one if the contemporaneous return response is negative (positive) during the day news announced. The estimation results are qualitatively similar to those we reported in Table 3.5 but lacks statistical significance. In another words, the statistical significance for positive news without attention and for positive news with retail attention disappears although the sign remains consistent with the results in Table 3.5. Considering that the magnitude on the coefficients for positive news with retail attention are also very similar between table 3.5 and table 3.8, the main take away for us is that using return sign as a proxy for news sentiment is consistent with directly

Table 3. 8: The Effect of Negative vs Pure Positive Sentiment Conditioning on Attention Types – Alternative Measure of Sentiment

This table provides estimation results of panel regression analysis for equation (3.6) in the sample period between Jan 1st, 2014 and Dec 31st, 2018, using alternative measure of sentiment. The dependent variables are DGTW-adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Negative (Positive) is an indicator variable equals one if the contemporaneous market return is negative (positive) during the day news announced. INST is an indicator variable equals one if there is only institutional attention and zero otherwise. RETL is an indicator variable equals one if there is only retail attention and zero otherwise. BOTH is an indicator variable equals one if there are both institutional and retail attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are clustered at firm and date level. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	[-1,0]	**[1, 5]**	**[1, 10]**	**[1, 15]**	**[1, 20]**
Negative	-0.519***	0.065	0.115	0.139	0.182
	(-15.31)	(1.64)	(1.60)	(1.41)	(1.52)
Negative*INST	-0.523***	-0.103*	-0.186**	-0.171*	-0.238*
	(-7.74)	(-1.69)	(-2.23)	(-1.66)	(-1.95)
Negative*RETL	-0.206	-0.073	-0.100	-0.199	-0.231
	(-1.28)	(-0.48)	(-0.52)	(-0.79)	(-0.80)
Negative*BOTH	-2.445***	-0.021	-0.071	-0.113	-0.228
	(-9.84)	(-0.16)	(-0.38)	(-0.52)	(-0.94)
Positive	0.506***	0.032	0.051	0.061	0.093
	(17.07)	(0.88)	(0.74)	(0.62)	(0.77)
Positive*INST	0.504***	0.040	0.041	0.079	0.008
	(7.58)	(0.64)	(0.48)	(0.73)	(0.06)
Positive*RETL	0.482***	0.041	-0.048	-0.324	-0.439
	(3.27)	(0.33)	(-0.27)	(-1.40)	(-1.57)
Positive*BOTH	2.592***	0.135	0.029	0.157	0.156
	(10.21)	(0.72)	(0.15)	(0.73)	(0.69)
INST	0.066**	-0.002	0.054	0.059	0.071
	(2.28)	(-0.05)	(1.00)	(0.82)	(0.89)
RETL	0.109***	0.050	0.111	0.108	0.148
	(2.73)	(0.99)	(1.36)	(1.19)	(1.38)
BOTH	0.030	-0.070	-0.059	-0.161	-0.059
	(0.28)	(-0.67)	(-0.41)	(-1.04)	(-0.34)
Observations	219,487	219,487	219,487	219,487	219,487
Adjusted R-squared	0.133	0.002	0.003	0.003	0.004
Controls	Yes	Yes	Yes	Yes	Yes
DATE FE	Yes	Yes	Yes	Yes	Yes

measuring sentiment of news through textual based approach with machine learning technics used in our study, but perhaps unsurprisingly, appears to be a noisy measure. This finding reinforces our choice of using textual-based approach to measure sentiment as it provides a more accurate measure.

3.5.2 Fama MacBeth Approach

We would also perform multivariate analysis using the Fama and MacBeth (1973) method. In the univariate analysis, we applied a two-stage approach which is similar to the approach in Fama and MacBeth (1973).

Table 3.9 provides the estimation results using the specification outlined in equation (3.6) by using Fama and MacBeth approach. The results are consistent with what found using panel regression approach. We find that both the magnitude and statistical significance are similar to those we reported in table 3.5. This suggests that our results are insensitive to the method we applied for the estimation.

Table 3. 9: The Effect of Negative vs Pure Positive Sentiment Conditioning on Attention Types estimated using Fama MacBeth Approach

This table provides estimation results of equation (3.6) in the sample period between Jan 1st, 2014 and Dec 31st, 2018 using Fama Macbeth (1973) approach. The dependent variables are DGTW adjusted cumulative abnormal returns in the window of [-1, 0], [1, 5], [1, 10], [1, 15], and [1, 20]. Negative is an indicator variable equals one if there is negative news released and zero otherwise. Pure Positive is an indicator variable equals one if there is only positive news announced for the same firm in the same date. INST is an indicator variable equals one if there is only institutional attention and zero otherwise. RETL is an indicator variable equals one if there is only retail attention and zero otherwise. BOTH is an indicator variable equals one if there are both institutional and retail attention and zero otherwise. Controls contain a broad range of firm characteristic variables such as log of market equity (ln(ME)), book-to-market ratio (BM), daily abnormal volume (AVOL), standard deviation of daily returns (SDRET), (high – low)/high (HLtH), daily bid-ask spread (Spread), daily stock turnover (Turnover), and lagged return (CAR[-5,-2]). The standard errors are Newey West HAC standard errors. *, **, *** indicates significance at 10%, 5%, and 1% level.

	DGTW-Adjusted Cumulative Abnormal Return				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	[-1,0]	**[1, 5]**	**[1, 10]**	**[1, 15]**	**[1, 20]**
Negative	-0.087***	0.058	0.038	0.019	0.054
	(-3.52)	(1.30)	(0.60)	(0.26)	(0.62)
Negative*INST	-0.174***	-0.128*	-0.087	0.038	-0.009
	(-3.52)	(-1.80)	(-0.86)	(0.30)	(-0.06)
Negative*RETL	-0.524**	-0.060	0.071	-0.496	-0.292
	(-2.31)	(-0.20)	(0.19)	(-1.09)	(-0.58)
Negative*BOTH	-0.516**	-0.036	-0.108	-0.053	0.160
	(-2.32)	(-0.16)	(-0.37)	(-0.17)	(0.45)
Pure Positive	0.011	0.037	0.099**	0.185***	0.245***
	(0.66)	(1.24)	(2.31)	(3.64)	(4.24)
Pure Positive*INST	0.283***	-0.009	-0.079	-0.080	-0.106
	(6.51)	(-0.14)	(-0.84)	(-0.70)	(-0.78)
Pure Positive*RETL	0.441***	-0.029	-0.195	-0.535**	-0.597*
	(2.62)	(-0.18)	(-0.87)	(-1.96)	(-1.86)
Pure Positive*BOTH	1.229***	0.314	0.188	0.184	0.000
	(5.42)	(1.61)	(0.72)	(0.58)	(0.00)
INST	0.042*	0.000	0.091*	0.059	0.056
	(1.91)	(0.00)	(1.69)	(0.97)	(0.81)
RETL	0.070**	0.014	0.048	0.093	0.118
	(2.36)	(0.27)	(0.63)	(1.00)	(1.11)
BOTH	-0.007	-0.141	-0.121	-0.188	-0.036
	(-0.06)	(-1.12)	(-0.77)	(-1.05)	(-0.18)
Observations	1213	1213	1213	1213	1213
Controls	Yes	Yes	Yes	Yes	Yes

3. 6 Conclusion

There is a debate in the literature on the role played by different types of investor attention for achieving market efficiency, and conflicting results for the impact of investor attention. While institutional attention is typically found to be stabilizing, some work demonstrates that retail attention is destabilizing and some work finds it stabilizing to financial markets. This paper contributes to the debate by considering the type of media news (complex versus simple), the sentiment of that news (positive versus negative versus mixed or neutral), and investor attention to news (institutional versus local retail versus national retail). Work in the field has typically considered proxies for news rather than direct measures as we do, typically ignores news complexity, often obscures the impact of retail investors on market prices and when retail investors are considered, only a very few papers have disentangled informed retail investors from the uninformed.

We show that retail attention does indeed destabilize financial markets by inducing price overreactions to positive news, but only if it is from uninformed retail investors. We find that when retail attention destabilizes the market it is when retail investors appear to struggle digesting complex business information and then only if the news is of a positive sentiment; negative sentiment news and retail investor attention are not associated with market instability, likely an outcome of retail investor's well-documented reluctance to short sell on negative news. This also likely reflects the retail investors' incapacity to correctly interpret the usefulness of the complicated information as well as their tendency of overreacting to attention-grabbing positive media news. Studies that do not find statistical significance for the destabilizing role of retail attention likely obscure this by mixing positive and negative news events together.

We find that institutional attention plays a stabilizing role in any context we explore, complex or simple news, positive or negative news sentiment, with or without retail investor attention, consistent with institutional investors being the smart money. Studies that find investor attention is associated with market instability are almost surely obscuring the stabilizing effect of institutional attention if they do not separate out news events to which only retail investors are attentive. We also find that the price overreaction induced by retail attention is only apparent when institutional attention is absent. Studies that fail to find retail attention to be destabilizing likely do not focus on news which institutional investors ignore.

Our exploration of the role of a specific type of retail attention – local retail attention, is particularly important. These are retail investors possessing an informational advantage for local firms. We find this subgroup of retail investors appear to be smart investors and appear to improve market efficiency (prices move with their attention to news and do not subsequently drift or reverse). This may explain why previous studies show mixed results on the role played by retail investors. Retail investors are a complicated set of investors who can, in the right context, stabilize markets for some firms (local to these investors) and simultaneously destabilize financial markets for other firms (that they are not local to). The aggregated effects of retail attention is therefore context specific as to which subgroup of retail investors play the dominant role.

REFERENCES

Agarwal, A., Leung, A. C.M., Konana, P., and Kumar, A., 2017. “Cosearch attention and stock return predictability in supply chains.” Information Systems Research 28(2): 265-288.

Ahern, K.R., and Sosyura, D., 2014. “Who writes the news? Corporate press releases during merger negotiations”, Journal of Finance 69(1): 241-291.

Alok, S., Kumar, N., and Wermers, R. 2020. “Do fund managers misestimate climatic disaster risk.” Review of Financial Studies, 33(3): 1146-1183.

Andrei, D., Cujean, J. 2017. “Information percolation, momentum and reversal.” Journal of Financial Economics, 123(3): 617-645.

Anderson, A. G., and Larkin, Y., 2018, “Does non-informative text affect investor behavior?”, Financial Management, forthcoming.

Babolmorad, N., and Massoud, N., 2020. “When sentiment is news: Topic-Adaptive Syntax Approach (TASA)”, working paper, Melbourne Business School, University of Melbourne.

Baker, S. R., Bloom, N., and Davis, S. J., 2016. “Measuring economic policy uncertainty.” The Quarterly Journal of Economics, 131(4): 1593-1636.

Bali, T. G., Cakici N., and Whitelaw R., 2011. “Maxing out: stocks as lotteries and the cross-section of expected returns.” Journal of Financial Economics, 99(2): 427–446.

Bali, T. G., Hirshleifer, D. A., Peng, L., and Tang, Y., 2019. “Attention, social interaction, and investor attraction to lottery stocks.” Working Paper, 9th Miami Behavioral Finance Conference 2018.

Barber B. M., Huang X., and Odean T., 2016. “Which factors matter to investors? Evidence from mutual fund flow.” The Review of Financial Studies, 29(10): 2600–2642.

Barber, B. M., and Odean, T., 2000. “Trading is hazardous to your wealth: The common stock investment performance of individual investors.” Journal of Finance 55(2): 773-806.

Barber, B. M., and Odean, T., 2008. “All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors.” Review of Financial Studies 21(2): 785-818.

Barber, B. M., Odean, T., and Zhu, N. 2009. “Do retail trades move markets?” The Review of Financial Studies 22(1): 151-186.

Banerjee, S., Dasgupta, S. and Kim, Y., 2008. “Buyer–supplier relationships and the stakeholder theory of capital structure.” Journal of Finance 63(5): 2507-2552.

Basistha, A., Kurov, A., and Wolfe, M.H., 2019. "Volatility forecasting: The role of internet search activity and implied volatility.", Available at SSRN: https://ssrn.com/abstract=2812387 or http://dx.doi.org/10.2139/ssrn.2812387.

Ben-Rephael, A., Da, Z., and Israelsen, R.D., 2017. "It depends on where you search: institutional investor attention and underreaction to news." The Review of Financial Studies 30(9): 3009-3047.

Ben-Rephael, A., Carlin, B.I., Da, Z., and Israelsen R.D., 2020. "Information consumption and asset pricing." Journal of Finance, forthcoming.

Ben-Rephael, A., Da, Z., Easton, P., and Israelsen R.D., 2018. "Who pays attention to SEC form 8-K?" Working Paper, Indiana University.

Berk, J.B., and van Binsbergen, J.H., 2015. "Measuring skill in the mutual fund industry." Journal of Financial Economics, 118(1): 1-20.

Bernile, G., Bhagwat, V., Kecskes, A., and Nguyen, P. A., 2018. "Are the risk attitudes of professional investors affected by personal catastrophic experiences?" Available at SSRN 3024983. Working Paper, University of Miami, George Washington University, York University.

Bernile, G., Kumar, A., and Sulaeman, J., 2015. "Home away from home: geography of information and local investors." Review of Financial Studies 28(7): 2009-2049.

Birru, J., Chague, F., De-Losso, R., and Giovannetti, B., 2020. "Attention and biases: Evidence from tax-inattentive investors." Fisher College of Business Working Paper No. 2019-03-022, Charles A. Dice Center Working Paper No. 2019-22, Available at SSRN: https://ssrn.com/abstract=3459478 or http://dx.doi.org/10.2139/ssrn.3459478

Boehmer, E., Jones, C. M., Zhang, X., and Zhang, X., 2020. "Tracking retail investor activity." Journal of Finance, Forthcoming, Available at SSRN: https://ssrn.com/abstract=2822105 or http://dx.doi.org/10.2139/ssrn.2822105.

Bordalo, P., Gennaioli, N., and Shleifer, A., 2012. "Salience theory of choice under risk." The Quarterly Journal of Economics, 127(3): 1243-1285.

Burt, A., and Hrdlicka, C.M., 2021. "Where does the predictability from sorting on returns of economically linked firms come from?" Journal of Financial and Quantitative Analysis 56(8): 2634-2658..

Carhart, M.M., 1997. "On persistence in mutual fund performance." Journal of Finance 52(1): 57-82.

Cen, L., Hertzel, M., and Schiller, C.M., 2019. "Speed matters: supply-chain information diffusion and Price feedback effects." Working Paper, University of Toronto. Available at SSRN: https://ssrn.com/abstract=3430225 or http://dx.doi.org/10.2139/ssrn.3430225

Cen, L., Danesh, E., Ornthanalai, C., and Zhao, X., 2015. “The power of economic networks: investor recognition through Supply-Chain Relationships.” Working Paper, University of Toronto.

Cohen, L. , Frazzini, A. , and Malloy, C.J., 2008. “The small world of investing: board connections and mutual fund returns.” Journal of Political Economy 116(5): 951–979.

Cohen, L., and Frazzini, A., 2008. “Economic links and predictable returns.” Journal of Finance 63(4): 1977-2011.

Chuprinin, O., and Gorbenko, A. and Kang, C. M., 2019, “Rationally neglected stocks”. Available at SSRN: https://ssrn.com/abstract=3285757 or http://dx.doi.org/10.2139/ssrn.3285757.

Coval, J.D. , and Moskowitz, T.J., 1999. “Home bias at home: local equity preference in domestic portfolios.” Journal of Finance 54(6), 2045–2073.

Coval, J. D., and Moskowitz, T. J., 2001. “The geography of investment: Informed trading and asset prices.” Journal of political Economy, 109(4): 811-841.

Cziraki, P., Mondria, J., and Wu, T., 2019. “Asymmetric attention and stock returns.” Management Science, Forthcoming.

Da, Z., Engelberg, J., and Gao, P., 2011. “In search of attention.” Journal of Finance 66(5): 1461 – 1499.

Da, Z., Hua, J., Hung, C., and Peng, L., 2020. “Market returns and a tale of two attentions.” Available at SSRN: https://ssrn.com/abstract=3551662 or http://dx.doi.org/10.2139/ssrn.3551662.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R., 1997. “Measuring mutual fund performance with characteristic-based benchmarks”, Journal of Finance 52(3): 1035-1058.

De Long, J. B., Shleifer A., Summers L. H., and Waldmann R. J., 1990. “Noise trader risk in financial markets.” Journal of Political Economy, 98(4): 703-738.

DellaVigna, S., and Pollet, J.M., 2009. “Investor inattention and Friday earnings announcements.” Journal of Finance 64(2): 709-749.

Dessaint, O., and Matray, A., 2017. “Do managers overreact to salient risks? Evidence from hurricane strikes.” Journal of Financial Economics, 126(1): 97-121.

Drake, M., Roulstone, D., and Thornock, J., 2015. “The determinants and consequences of information acquisition via EDGAR.” Contemporary Accounting Research 32(3): 1128-61.

Duffie, D., Malamud, S., and Manso, G., 2009. “Information percolation with equilibrium search dynamics.” Econometrica, 77(5): 1513-1574.

Engelberg, J.E., Reed, A.V., and Ringgenberg, M.C., 2012. “How are shorts informed?: Short sellers, news, and information processing”, Journal of Financial Economics 105(2), 260-278.

Eraker, B., and Ready, M., 2015. “Do investors overpay for stocks with lottery-like payoffs? An examination of the returns of OTC stocks.” Journal of Financial Economics, 115(3): 486-504.

Fama, E.F., and MacBeth, J.D., 1973. “Risk return and equilibrium: Empirical tests.” Journal of Political Economy 81(3), 607–636.

Fama, E.F., and French , K.R., 1993. “Common risk factors in the returns on stocks and bonds.” Journal of Financial Economics 33(1): 3-56.

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

Fedyk, A., 2021. “Disagreement after news: Gradual information diffusion or differences of opinion?”, Review of Asset Pricing Studies 11(3):165-501.

Frank, M.Z., and Sanati, A., 2018. “How does the stock market absorb shocks?”, Journal of Financial Economics, 129(1): 136-153.

French, K.R., 2008. “Presidential address: The cost of active investing.” Journal of Finance 63(4): 1537-73.

Gacia, D., 2013. “Sentiment during recessions”, Journal of Finance 68(3): 1267-1300.

Galai, D., and Sade, O., 2006. “The ‘ostrich effect’ and the relationship between the liquidity and the yields of financial assets.” Journal of Business, 79(5), 2741-2759.

Gervais, S., Kaniel, R., and Mingelgrin, D.H., 2001. “The high-volume return premium.” Journal of Finance, 56(3): 877-919.

Guan, Y., Wong, M. F., and Zhang, Y., 2015. “Analyst following along the supply chain.” Review of Accounting Studies 20(1): 210-241.

Guiso, L., Sapienza, P., and Zingales, L., 2018. “Time varying risk aversion.” Journal of Financial Economics, 128(3): 403-421.

Han, B., Hirshleifer, D., and Walden, J., 2019. “Social transmission bias and investor behavior.” (No. w24281). National Bureau of Economic Research.

Hartzmark, S.M., 2014. “The worst, the best, ignoring all the rest: The rank effect and trading behavior”, Review of Financial Studies, 28(4): 1024-1059.

Hirshleifer, D., 2020. “Presidential address: Social transmission bias in economics and finance.” Journal of Finance, 75(4): 1779-1831.

Hirshleifer, D., Peng L., and Wang, Q., 2020. “Social networks and market reactions to earnings news.” Working Paper, University of California at Irvine, CUNY Baruch College, and Hong Kong Baptist University.

Hirshleifer, D., and Teoh, S.H., 2003. “Limited attention, information disclosure, and financial reporting. ” Journal of Accounting and Economics 36 (1): 337–86.

Hirshleifer, D., Hou, K., Teoh, S.H., and Zhang Y., 2004. “Do investors overvalue firms with bloated balance sheets?” Journal of Accounting and Economics 38(1): 297-331.

Hirshleifer, D., Lim, S.S., and Teoh, S.H., 2011. “Limited investor attention and stock market misreactions to accounting information.” Review of Asset Pricing Studies 1(1): 35-73.

Hong, H. , Kubik, J.D. , and Stein, J.C., 2005. “Thy neighbor’s portfolio: word-of–mouth effects in the holdings and trades of money managers.” Journal of Finance 60(6): 2801–2824.

Hong, H., Kubik, J.D., and Stein, J.C., 2008. “The only game in town: Stock-price consequences of local bias.” Journal of Financial Economics 90(1): 20-37.

Hong, H., and Stein, J.C., 1999. “A unified theory of underreaction, momentum trading, and overreaction in asset markets.” Journal of Finance 54(6): 2143-2184.

Hong, H. G., and Xu, J., 2019. “Inferring latent social networks from stock holdings.” Journal of Financial Economics, 131(2): 323-344,

Hou, K., 2007. “Industry information diffusion and the lead-lag effect in stock returns.” The Review of Financial Studies 20(4): 1113-1138.

Hou, K., Peng, L., and Xiong, W., 2009. “A tale of two anomalies: the implications of investor attention for price and earnings momentum.” Working Paper, Available at SSRN: https://ssrn.com/abstract=976394.

Huberman, G., 2001. “Familiarity breeds investment.” The Review of Financial Studies, 14(3): 659-680.

Hvidkjaer, S. 2008. “Small trades and the cross-section of stock returns.” The Review of Financial Studies 21(3): 1123-1151.

Ivković, Z., and Weisbenner, S., 2005. “Local does as local is: Information content of the geography of individual investors' common stock investments.” Journal of Finance 60(1): 267-306.

Jegadeesh, N., 1990. “Evidence of predictable behavior of security returns.” Journal of Finance 45(3): 881-898.

Jegadeesh, N., and Titman, S., 1993. “Returns to buying winners and selling losers: Implications for stock market efficiency,” Journal of Finance 48(1), 65–91.

Jiang, H., and Verardo, M., 2018. “Does herding behavior reveal skill? An analysis of mutual fund performance.” The Journal of Finance, 73(5): 2229-2269.

Kahneman, Daniel. 1973. “Attention and effort.” Englewood Cliffs, N.J.: Prentice-Hall.

Kaniel, R., Saar, G., and Titman, S. 2008. “Individual investor trading and stock returns.” Journal of Finance 63(1): 273-310.

Kaniel, R., Liu, S., Saar, G., and Titman, S., 2012. “Individual investor trading and return patterns around earnings announcements.” Journal of Finance 67(2): 639-680.

Kaplanski, G., and Levy, H., 2010. “Sentiment and stock prices: The case of aviation disasters.” Journal of Financial Economics, 95(2): 174-201.

Kelly, E. K., and Tetlock, P. C., 2013. “How wise are crowds? Insights from retail orders and stock returns.” Journal of Finance 68(3): 1229-1265.

Kelly, E. K., and Tetlock, P. C., 2017. “Retail short selling and stock prices.” The Review of Financial Studies 30(3): 801-834.

Kuchler, T., Li Y., Peng L., Stroebel J., and Zhou D., 2020. “Social proximity to capital: Implications for investors and firms.” Working Paper, New York University and Baruch College.

Kumar, A., and Lee, C. M.C., 2006. “Retail investor sentiment and return comovements.” Journal of Finance 61(5): 2451-2486.

Loughran, T., and McDonald, B., 2011. “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks”, Journal of Finance 66(1): 35-65.

Lee, C. M.C., Sun, S.T., Wang, R., and Zhang, R., 2019. “Technological links and predictable returns.” Journal of Financial Economics 132(3): 76-96.

Liu, H., Peng, L., and Tang, Y., 2019. “Retail attention, institutional attention.” Baruch College Zicklin School of Business Research Paper No. 2019-05-03, Available at SSRN: https://ssrn.com/abstract=3370683 or http://dx.doi.org/10.2139/ssrn.3370683.

Ma, J., Xiong, X., and Feng, X., 2020. “News release and the role of different types of investors”, International Review of Financial Analysis 73(2): 101643.

Massa, M. and Simonov, A., 2006. “Hedging, familiarity and portfolio choice.” The Review of Financial Studies 19(2): 633-685.

Menzly, L., and Ozbas, O., 2010. “Market segmentation and cross-predictability of returns.” Journal of Finance 65(4): 1555-1580.

Merton, R., 1987. “A simple model of capital market equilibrium with incomplete information.” Journal of Finance 42(3): 483-510.

Moskowitz, T., and Grinblatt, M., 1999. “Do industries explain momentum?,” Journal of Finance 54(4), 1249–1290.

Pool, V. K., Stoffman, N., and Yonker, S. E., 2012. “No place like home: Familiarity in mutual fund manager portfolio choice.” The Review of Financial Studies, 25(8): 2563-2599.

Pástor, L., and Stambaugh, R.F., 2003. “Liquidity risk and expected stock returns." Journal of Political Economy 111(3): 642-685.

Peng, L., and Xiong, W., 2006. “Investor attention, overconfidence and category learning,” Journal of Financial Economics 80(3): 563-602.

Pritamani, M., and Singal, V., 2001. “Return predictability following large price changes and information releases”, Journal of Banking and Finance, 25(4): 631-656.

Rhodes–Kropf, M., Robinson, D. T., and Viswanathan, S., 2005. “Valuation waves and merger activity: The empirical evidence.” Journal of Financial Economics 77(3): 561-603.

Savor, P.G., 2012. “Stock returns after major price shocks: The impact of information”, Journal of Financial Economics, 106(3): 635-659.

Seasholes, M.S., and Zhu, N., 2013. “Investing in what you know: The case of individual investors and local stocks.”, Journal of Investment Management, 11(1): 20-30.

Shefrin, H., and Stataman, M., 1985. “The disposition to sell winners too early and ride losers too long: theory and evidence.” Journal of Finance, 40(3): 777-790.

Stambaugh, R.F., 2014. “Presidential address: investment noise and trends,” Journal of Finance 69(4): 1415-1453.

Soloman, D.H., 2012. “Selective Publicity and Stock Prices”, Journal of Finance 67(2): 599-638.

Song, S., 2020. “Does individual investor attention to accounting information influence the pricing of stocks?”. INSEAD Working Paper No. 2020/15/ACC, Available at SSRN: https://ssrn.com/abstract=3229526 or http://dx.doi.org/10.2139/ssrn.3229526.

Tetlock, P.C., 2007. “Giving content to investor sentiment: The role of media in the stock market”, Journal of Finance, 62(3): 1139-1168.

Tetlock, P.C., 2014. “All the news that’s fit to reprint: Do investors react to stale information?”, Review of Financial Studies, 24(5): 1481-1512.

Tetlock, P.C., Saar-Tsechansky, M., and Macskassy, S., 2008. “More than words: quantifying language to measure firm’s fundamentals”, Journal of Finance 63(3): 1437-1467.

Wermers, R., 2004. “Is money really ‘smart’? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence”, Working Paper, Available at SSRN: https://ssrn.com/abstract=414420 or http://dx.doi.org/10.2139/ssrn.414420.

Zhou, X., 2020. “Whose attention matters? Evidence from the return predictability between economically linked firms”, Working Papers, Schulich School of Business, York University.