Journal of Behavioral and Experimental Finance 39 (2023) 100820

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Finance

journal homepage: www.elsevier.com/locate/jbef

Full length article

Active attention, retail investor base, and stock returns

Zhongdong Chen [a,*], Karen Ann Craig [b]

[a] Department of Finance, College of Business, University of Northern Iowa, 309 Curris Business Building, Cedar Falls, IA 50614-0129, United States of America

[b] Department of Accounting, Finance and Information Systems, College of Business, Eastern Michigan University, 522 Hill Hall, Ypsilanti, MI 48197, United States of America

ARTICLE INFO

Article history:
Received 21 October 2022
Received in revised form 18 March 2023
Accepted 21 May 2023
Available online 29 May 2023

JEL classification:
G11
G14

Keywords:
Retail investor attention
Investor recognition
Google search volume index

ABSTRACT

The vast investor attention literature assumes that active retail investor attention drives retail investor base and therefore, stock demand and stock returns. We aim to provide empirical support for these critical assumptions that are heavily relied upon in the literature. Utilizing Robinhood investor data to measure retail investor base and the Google Search Volume Index for stock tickers to measure active retail investor attention, we find that active retail investor attention and increases in active retail investor attention are associated with a larger retail investor base, which is positively related to demand for stocks and stock returns in the subsequent four weeks. We find these effects are impacted by recent stock returns and are economically more significant for larger stocks.

1. Introduction

Employing measures of active retail investor attention, such as the Google Search Volume Index for stock tickers, a large number of empirical studies find that active retail investor attention is positively related to short-term stock trading volume and returns (Bae and Wang, 2012; Chen et al., 2021; Da et al., 2011; Ding and Hou, 2015; Kao et al., 2022; Kupfer and Schmidt, 2022; Rakowski et al., 2021; Swamy and Dharani, 2019; Takeda and Wakao, 2014; Yang et al., 2021; Ying et al., 2015; Zhang and Wang, 2015). In contrast to the mainstream asset pricing studies that focus on risk factors (Fama and French, 2018; Hou et al., 2015), these studies contend the investor recognition hypothesis explains their results,[1] theorizing that active retail investor attention is a proxy for stock visibility. This explanation relies on two critical assumptions pertaining to the investor recognition hypothesis: first, active attention positively affects retail investor base; and second, retail investor base positively impacts stock demand and stock returns. A similar argument is also found in the literature on bond valuation.[2] However, the empirical support for these two critical assumptions is limited to date, making them vulnerable to alternative views.[3] We aim to fill this gap in the broad investor attention literature and validate the two critical assumptions for the large number of studies on active retail investor attention.

The impact of attention on the trading behavior of retail investors is important to understand as retail investors continue to grow into a major force in the stock market. User friendly mobile trading platforms and reduced commission fees have resulted in skyrocketing retail investor participation in the U.S. stock market in recent years, comprising a significant and growing portion of market trades.[4] Claiming 14 million retail investor accounts in 2022 (Robinhood, 2022), Robinhood is one of the most popular retail trading platforms and its investor data are widely used

* Corresponding author.
E-mail addresses: zhongdong.chen@uni.edu (Z. Chen), kcraig9@emich.edu (K. Craig).

[1] Merton (1987) proposes the investor recognition hypothesis, which postulates that stocks with higher visibility have a larger investor base that favorably affects their demand and valuation.

[2] For example, Tang and Zhang (2020) argue that investor attention could potentially help expand the investor base for issuers of green bonds. Massa and Žaldokas (2014) hypothesize that attracting investor attention will increase the demand of new investors for international bonds issued by U.S. firms.

[3] For example, active retail investor attention and stock demand measures (e.g., the number of Google searches) may be driven by existing investor base, not new retail investors. In this case, these measures are not associated with a larger investor base, per the first critical assumption, or higher levels of stock visibility.

[4] In 2020, there were over 120 million retail investor accounts at the top five online brokerages including Fidelity, Vanguard, Schwab, Weibull, and Robinhood. In January 2020, approximately 17% of market trades were made by retail investors and six months later, this number increased to more than 25%. (McCrank, 2021).

https://doi.org/10.1016/j.jbef.2023.100820
2214-6350/

in studies on retail investor behavior (Barber et al., 2022; Eaton et al., 2022; Liaukonyte and Zaldokas, 2022; Moss et al., 2020; Pagano et al., 2021; Welch, 2022).

We use Robinhood investor data to construct relative measures of weekly retail investor base to explore whether active attention positively affects retail investor base and if retail investor base positively impacts stock demand and stock returns. We utilize the Google Search Volume Index to measure active retail investor attention for a sample of 2,552 U.S. stocks in the period from May 2018 to August 2020, corresponding to the availability of Robinhood investor data.

The results show that (the change in) weekly active retail investor attention is positively related to (the change in) weekly retail investor base in the subsequent four weeks. We focus on the subsequent four weeks as the literature suggests a short-term impact of active retail investor attention on stocks (Da et al., 2011). This evidence indicates that active retail investor attention positively affects investor base, providing empirical support for the first critical assumption in the investor attention literature. More importantly, we find such relationship influenced by recent stock returns. With a higher level of active retail investor attention, winner stocks, with a positive weekly return, attract more retail investors than loser stocks, with a negative weekly return, in the subsequent four weeks. This result could be attributed to winner chasing of retail investors (Bailey et al., 2011; Haruvy et al., 2007). However, with a temporary increase in active retail investor attention, loser stocks tend to attract more retail investors than winner stocks in the subsequent weeks, likely driven by retail investors "buying the dip" (Dreman et al., 2001; Shohfi and Simaan, 2021). This new evidence suggests that active retail investor attention plays a significant role in the trading behavior of retail investors. These results, robust to alternative definitions of active retail investor attention and retail investor base, are not likely driven by the endogeneity between active retail investor attention and retail investor base.

We also find results supporting the second critical assumption in the investor attention literature. The size of retail investor base and the change in retail investor base are both positively related to demand for stocks measured with mean daily trading volume (Lo and Wang, 2006; Yezegel, 2015) in the subsequent 1-week, 2-week, 3-week, and 4-week windows. Consistent with this result, we find the size of retail investor base and the change in retail investor base both positively affect stock returns in the subsequent four windows, which is likely due to the price pressure created by higher demand associated with a larger retail investor base. This evidence supports the buying pressure hypothesis in the attention literature. In addition, with a larger retail investor base, winner stocks tend to have a lower return in the following four weeks, relative to loser stocks. With a temporary increase in retail investor base, winner stocks also tend to have a lower return in the following four weeks, relative to loser stocks (Jankensgård and Vilhelmsson, 2018). These results, robust to alternative definitions of retail investor base, could be explained by the disposition effect (Odean, 1998; Rau, 2015) where retail investors tend to sell winner stocks and hold loser stocks.

Overall, our results show that active retail investor attention positively impacts retail investor base which, in turn, affects demand for stocks and stock returns. However, in contrast to the conjecture in the previous studies, results indicate that active retail investor attention has an economically more significant impact on retail investor base of larger stocks than that of smaller stocks, in the subsequent four weeks. In line with those results, retail investor base has an economically more significant impact on larger stocks, in terms of stock demand and stock returns in the following four weeks, consistent with Meshcheryakov and Winters (2022). We also find retail investor base impacts annualized stock return volatilities in the subsequent weeks. Winner stocks with a larger retail investor base and winner stocks with a temporary increase in retail investor base tend to have higher annualized stock return volatilities in the following four weeks.

This study contributes to the vast investor attention studies (Barber and Odean, 2008; Chai et al., 2021; Israeli et al., 2022; Liaukonyte and Zaldokas, 2022; Lou, 2014; Madsen and Niessner, 2019; Mayer, 2021; Rakowski et al., 2021), especially those relying on measures of active retail investor attention such as the Google Search Volume Index. The results in this study validate the assumption in the literature that active retail investor attention has a positive impact on retail investor base, which in turn, positively impacts stock demand. These results contribute to the debate on the buying pressure hypothesis (Aouadi et al., 2013; Bank et al., 2011; Chen et al., 2021; Desagre and D'Hondt, 2021; Joseph et al., 2011). In addition, the new evidence, that the impacts of active retail investor attention and retail investor base are influenced by concurrent stock returns (e.g., winner chasing, buying the dip, and the disposition effect), links studies on active retail investor attention to those on the trading behavior of retail investors (Bailey et al., 2011; Birru, 2015; Haruvy et al., 2007; Liaukonyte and Zaldokas, 2022; Meng and Weng, 2018; Odean, 1998; Rau, 2015). Moreover, this study contributes to the strand of literature relating retail investors to stock return volatilities (Andrei and Hasler, 2015; Dimpfl and Jank, 2016; Hervé et al., 2019; Jankensgård and Vilhelmsson, 2018). Ultimately, this study supports empirical studies suggesting a positive relationship between Google search volume and the number of Google search users (Da et al., 2015; Kostopouls et al., 2020).

The remainder of this study is organized as follows: Section 2 presents the literature review, Section 3 describes the data and methodology, Section 4 reports the empirical analysis, Section 5 presents the robustness tests, and Section 6 concludes.

2. Literature review

2.1. Robinhood

Founded in 2013, Robinhood was the first fintech brokerage that offered commission-free trades in early 2015. Robinhood has both mobile- and web-based trading platforms for stock trades. The Robinhood simple and engaging mobile app is believed to be the most important reason for its success in attracting younger and less experienced retail investors. According to Robinhood's reports, the number of Robinhood net cumulative funded accounts is 3.3 million, 5.1 million, and 12.5 million in years 2018, 2019, and 2020, respectively. We therefore, include year fixed effects in the empirical tests to control for the increases in the number of Robinhood investors.

Robinhood investors are likely small and inexperienced investors (Barber et al., 2022; Eaton et al., 2022; Welch, 2022) and their trading behavior might be more heavily impacted by attention (Seasholes and Wu, 2007). Therefore, recent studies utilize Robinhood investor data to explore the impacts of retail investor attention on their trading behavior and the stock market (Barber et al., 2022; Eaton et al., 2022; Friedman and Zeng, 2021; Liaukonyte and Zaldokas, 2022; Moss et al., 2020; Pagano et al., 2021; Welch, 2022).

In addition to the convenient and user-friendly interface for stock trading, Robinhood also provides a function named "Browse" that aims to help investors discover new companies and stay informed about the market, popular companies, and daily winner/loser stocks. Although the information provided in "Browse" is not an investment recommendation (Robinhood), it has the potential to influence investors' attention. The "Browse" function would have an influence similar to any other online news or information forum that disseminates information. We use Google

Search Volume Index to measure retail investors' active attention from all external influences like the "Browse" function, because these external influences will likely motivate investors to use Google to find out more information.

2.2. Retail investor attention and the buying pressure hypothesis

Out of thousands of stocks, retail investors only hold a small fraction of available stocks due to the cost of gathering and analyzing the information and the cost of transmitting information between parties. As a result, retail investors only invest in stocks that have caught their attention (Peng and Xiong, 2006). Because of this, stocks that can attract the attention of retail investors will experience a higher demand and subsequently, higher returns. Recent studies have found empirical results consistent with this buying pressure hypothesis (Bank et al., 2011; Joseph et al., 2011; Takeda and Wakao, 2014; Vlastakis and Markellos, 2012; Aouadi et al., 2013). For instance, Barber and Odean (2008) find that retail investors are the net buyers of attention-grabbing stocks, resulting in buying pressure for those stocks. Da et al. (2011) find a positive relationship between retail investor attention and higher stock returns on Russell 3000 stocks and attribute this result to buying pressure. Chen et al. (2021) find that depressed stocks at year end are more likely to attract the attention of risk-seeking retail investors and experience increases in demand, leading to higher buying pressure and higher returns of the depressed stocks in early January.

However, examining trading accounts of Belgian retail investors, Desagre and D'Hondt (2021) find that retail investor attention is positively related to both retail investor buying and selling, and this effect is not stronger for retail investor buying. This finding challenges the buying pressure hypothesis. In this study, we do not examine the impact of retail investor attention on the number of retail investors buying or selling a particular stock due to data availability. We focus on the impact of retail investor attention on the total number of retail investors for each stock and its influence on subsequent stock demand and stock returns.

2.3. Passive measures of retail investor attention

Barber and Odean (2008) argue that passive attention measures including exposure in social media, high trading volume, and extreme daily returns can catch the attention of retail investors, making them net buyers of attention-grabbing stocks (Barber and Odean, 2008; Chen, 2021; Dang et al., 2018; Ding et al., 2020; Gervais et al., 2001; Kaniel et al., 2012; Liaukonyte and Zaldokas, 2022; Madsen and Niessner, 2019; Rakowski et al., 2021). In addition, advertising can also attract retail investors. Prior studies find that advertising expenses positively affect retail investor base, the buying of retail investors, and short-term stock returns (Chemmanur and Yan, 2019; Grullon et al., 2004; Lou, 2014; Mayer, 2021).

In contrast, we employ weekly active measures of retail investor attention constructed with the Google Search Volume Index, rather than passive measures such as quarterly or annual advertising expenses. As discussed in the following section, using active measures of retail investor attention ensures a direct and timely determinate of retail investors' interest in a stock.

2.4. Active measures of retail investor attention

The Google Search Volume Index for stock tickers is considered an active, direct, and timely measure of retail investor attention for stocks (Da et al., 2011; Desagre and D'Hondt, 2021). It measures the frequency a ticker is searched using Google during a period, scaled by the maximum frequency over the entire sampling horizon, giving the index a value between 0 and 100.

Prior studies find the Google Search Volume Index for stock tickers predictive of short-term trading volume and stock returns. Da et al. (2011) find a positive relationship between the Google Search Volume Index for tickers and stock returns in the coming weeks. Chen et al. (2021) find that increases in the Google Search Volume Index for tickers at year end predict higher trading volumes and higher stock returns in early January. Similarly, Takeda and Wakao (2014), Swamy and Dharani (2019), and Joseph et al. (2011) also find the Google Search Volume Index for tickers is related to trading volume and stock returns. In addition, studies using other search indexes for tickers find similar results (Yang et al., 2021; Ying et al., 2015; Zhang and Wang, 2015).

These studies argue that the Google Search Volume Index for tickers positively impacts trading volume and stock returns because this active attention measure is positively related to the visibility of stocks and consequently, leads to a larger retail investor base. Yet there is little empirical evidence to support the assumed relationship between the Google Search Volume Index for tickers and retail investor base, and it is possible that Google searches are driven by existing retail investors instead of new investors. We aim to fill this gap in the literature. The only prior empirical study linking the Google Search Volume Index for tickers to retail investor base is Ding and Hou (2015). However, there are important differences between this study and Ding and Hou (2015), making this study vital to provide stronger empirical support for the vast attention literature built on active measures of retail investor attention.

First, Ding and Hou (2015) utilize S&P 500 stocks for their sample evaluating retail investor attention and retail investor base. S&P 500 constituent stocks are the largest stocks with a minimum market capitalization of $8.20 billion in 2020 and positive cumulative earnings in the prior four quarters. This makes S&P 500 stocks highly selective and already highly visible to retail investors (Chen et al., 2004). In contrast, we employ a random sample of 2,552 unique U.S. stocks, enabling the comparison of the impacts of retail investor attention and retail investor base between smaller and larger stocks. Second, we measure retail investor base with weekly investor data from Robinhood while Ding and Hou (2015) examine annual investor base data from the Compustat database that includes all types of investors. The Robinhood investor data better represent the dynamic relationship between retail investor attention and retail investor base providing a clearer understanding of the impact of retail investor attention. Third, we examine the potential endogeneity between retail investor attention measured with the Google Search Volume Index for tickers and retail investor base, which is absent in the prior studies. Fourth, in contrast to prior studies that respectively focus on one of the two critical assumptions, we are able to test both assumptions within the same framework, providing holistic and stronger support for the literature. Finally, we examine the role of recent stock returns in the relationships among retail investor attention, retail investor base, and stock returns, which is absent in prior studies.

3. Data and methodology

3.1. Data and variables

Based on the available investor data from Robinhood (Robin-Track), we employ weekly observations in a 120-week sample period from May 2018 to August 2020.[5] We begin with stocks

[5] http://www.robintrack.net

Table 1
Summary statistics.

Variables	Mean	Median	25th Percentile	75th Percentile	Std. Dev.	n
ASVI	2.0782	2.4849	1.3863	3.0123	1.2143	277,650
Δ ASVI	0.0038	0.0000	−0.3137	0.3124	1.1315	277,650
GSVI	2.6416	3.2581	1.6094	3.8918	1.6097	277,650
Δ GSVI	0.0036	0.0000	−0.3075	0.3070	1.5003	277,650
Investors_Max	0.4179	0.3769	0.1940	0.6133	0.2856	277,650
Δ investors_Max	0.0056	0.0000	−0.0040	0.0065	0.0535	275,185
Investors_Ave	0.4060	0.3672	0.1880	0.5992	0.2641	277,650
Δ investors_Ave	0.0053	0.0004	−0.0033	0.0067	0.0308	275,185
Ret_1 W	0.0019	0.0043	−0.0133	0.0159	0.0344	277,650
Ret_1M	0.0073	0.0133	−0.0159	0.0396	0.0729	277,650
Ret_3M	0.0160	0.0119	−0.0304	0.0583	0.1217	277,650
Volatility	8.1218	6.5453	4.3462	9.9321	6.5357	277,650
Illiquidity	0.0290	0.0159	0.0040	0.0397	0.0373	277,650
Price	0.1844	0.0472	0.0197	0.1437	0.4517	277,650
BTM	0.5726	0.4273	0.1820	0.8055	1.9068	277,650
Size	13.7229	13.7034	12.1849	15.2303	2.2495	277,650
Institutional	0.5292	0.5690	0.2840	0.7291	0.3826	277,650
Analyst	1.7999	1.9459	1.0986	2.5649	0.9919	277,650
Buy	1.3374	1.3863	0.6931	2.0794	0.9132	277,650
Advertising	1.0498	0.0000	0.0000	1.4417	1.8126	277,650
Earnings	0.0256	0.0000	0.0000	0.0000	0.1579	277,650
Jump	0.6550	0.0000	0.0000	0.0000	0.2474	277,650
Volume	1.6391	0.9779	0.7344	1.3259	29.6563	277,650

This table reports summary statistics for 2,552 unique U.S. stocks in the May 2018–August 2020 sample period. *ASVI* is the average weekly Google Search Volume Index for "ticker", "ticker+stock", and "ticker+price". *GSVI* is the weekly Google Search Volume Index for "ticker". *ΔASVI* and *ΔGSVI* are the weekly changes in *ASVI* and *GSVI*, respectively. *Investors_Max* is the maximum number of retail investors for each stock in each week scaled by the maximum number of retail investors in the sample period. *Investors_Ave* is the average daily number of retail investors for each stock in each week scaled by the maximum number of retail investors in the sample period. *ΔInvestors_Max* and *ΔInvestors_Ave* are the weekly changes in *Investors_Max* and *Investors_Ave*, respectively. *Ret_1W*, *Ret_1M*, and *Ret_3M* are 1-week, 1-month, and 3-month stock returns adjusted with equal-weighted market returns, respectively. *Volatility* is annualized stock return volatility calculated using daily stock returns in the prior year. *Illiquidity* is zero return day ratio calculated as the number of zero return trading days divided by total number of trading days in the prior year. *Price* is price level measured with one divided by stock price. *BTM* is book-to-market ratio and *Size* is the natural logarithm of market capitalization. *Institutional* is institutional ownership. *Analyst* and *Buy* are the natural logarithms of the number of analysts following and the number of buy recommendations, respectively. *Advertising* is the natural logarithm of annual advertising expenses in thousands. *Earnings* is an earnings announcement indicator that is equal to one for weeks with an earnings announcement and zero otherwise. *Jump* is an indicator that is equal to one for weeks with a stock price jump and zero otherwise. *Volume* is the ratio between mean daily trading volume in a week and that in the prior week.

covered in both the Center for Research in Security Prices (CRSP) database and the Robinhood investor database, and collect the weekly Google Search Volume Index data for their tickers to measure weekly retail investor attention for those stocks.[6] The challenge of using the Google Search Volume Index for tickers to measure retail investor attention is that some tickers might have other significant meanings and therefore, the Google Search Volume Index data for those tickers may not accurately reflect retail investors' interest in the associated stocks. Therefore, we exclude tickers with a generic meaning (Da et al., 2011). We individually search for the remaining tickers on Google and exclude tickers that do not appear on the first page of the Google search results as Google presents the most relevant search results first (Chen et al., 2021). This selection process yields a sample of 2,552 unique stocks and 277,650 stock-week observations in the sample period.

Following the literature (Chen et al., 2021; Da et al., 2011; Meshcheryakov and Winters, 2022), we employ two measures for retail investor attention. The first is the weekly Google Search Volume Index for tickers (GSVI). The second is the weekly average Google Search Volume Index of three search terms (ASVI), including "ticker", "ticker+stock", and "ticker+price".[7] This second measure further addresses the issue where tickers might have other significant meanings and it also captures retail investor groups with different search habits. We follow the literature and use the natural logarithm of the two measures in the empirical tests. Summary statistics for these two measures and the weekly changes in these measures are reported in Table 1. We report results for ASVI in the main part of the paper and results for GSVI in robustness tests.

RobinTrack reports the daily number of total Robinhood investors holding a particular stock.[8] We use this data to construct two weekly measures for retail investor base. The first we dub the maximum retail investor base, which is the maximum number of Robinhood investors for each stock in each week. The second is the average retail investor base, determined by the average daily number of Robinhood investors for each stock in each week. We scale these two measures with the maximum number of Robinhood investors of each stock in the entire sample period so that our results will not be biased by stocks with a small number of investors at the beginning of the sample period. Additionally, our main independent variable of interest, Google Search Volume Index, is also a relative measure constructed in a similar manner. We report results for average retail investor base in the main part of the paper and results for maximum retail investor base in robustness tests.

Data for the control variables are obtained from the CRSP, Compustat, and Bloomberg databases. We control for 1-week, 1-month, and 3-month stock returns adjusted with equal-weighted

[6] Google Trends provides daily Google Search Volume Index data only for sample periods shorter than nine months.

[7] Arguably, the Google Search Volume Index for "ticker+stock" and "ticker+price" might be better able to capture retail investor attention on associated stocks. But similar to the findings in Chen et al. (2021), the Google Search Volume Index values for those two search terms are low relative to that for "ticker" and for many stocks, and the values are zero in a large number of weeks.

[8] RobinTrack does not report the daily number of new Robinhood investors that have purchased a particular stock or the daily number of existing Robinhood investors that have sold a particular stock. The daily number of total Robinhood investors reported by RobinTrack is the net result of existing Robinhood investors selling and new Robinhood investors buying. Therefore, we are not able to examine the impact of retail investor attention on the number of retail investors buying or selling a particular stock. In addition, RobinTrack does not report the number of shares Robinhood investors hold.

market returns in the empirical tests to capture the momentum and contrarian trading strategies of retail investors (Kaniel et al., 2008; Solomon et al., 2014).[9] In addition, we control for annualized stock return volatilities, illiquidity, price level, book-to-market ratio, and size. Annualized stock return volatilities are calculated using daily returns in the prior year, illiquidity is measured by the zero return day ratio in the prior year calculated as the ratio between the number of zero return trading days and the total number of trading days (Lesmond et al., 1999), price level is measured with one divided by stock price (Poterba and Weisbenner, 2001) in the week where attention is measured, book-to-market ratio is calculated as book value per share in the prior year divided by stock price in the week where attention is measured, and size is proxied by the natural logarithm of market capitalization in the week where attention is measured.

Institutional ownership in the prior quarter, the natural logarithm of the number of analysts following in the prior quarter, the natural logarithm of the number of buy recommendations in the prior quarter, and the natural logarithm of annual advertising expenses in the prior year[10] are also controlled in the empirical tests as they might impact the attention of retail investors (Chen et al., 2021; Lee and So, 2017; Lou, 2014; Kumar and Lee, 2006; Madsen and Niessner, 2019). Moreover, we control for earning announcements in the week where attention is measured, with an indicator equal to one for weeks with an earnings announcement and zero otherwise, as earnings announcements might attract retail investors (Chen et al., 2021).

Ultimately, we control for information shocks and volume shocks, which might catch the attention of retail investors. Following Jiang and Zhu (2017) and others, we employ jumps in stock price in the week where attention is measured, identified using the swap variance approach (Jiang and Oomen, 2008), as a proxy for information shocks. This indicator variable is equal to one for weeks with a price jump, and zero otherwise. Volume shocks are measured with the ratio between mean daily trading volume in the week where attention is measured and that in the prior week.[11]

3.2. The models

To test the first critical assumption in the vast literature on active measures of retail investor attention, we estimate OLS Model (1) to examine whether our active measures of retail investor attention are positively associated with retail investor base:

$$\begin{aligned} Investors_{i,t,q,n} &= \beta_0 + \beta_1 Attention_{i,t-1} + \beta_2 Attention_{i,t-1} \\ &* Ret\_1W_{i,t-1} + \beta_3 Ret\_1W_{i,t-1} + \beta_4 Ret\_1M_{i,t-1} + \beta_5 Ret\_3M_{i,t-1} \\ &+ \beta_6 Volatility_{i,n-1} + \beta_7 Illiquidity_{i,n-1} + \beta_8 Price_{i,t-1} + \beta_9 BTM_{i,t-1} \\ &+ \beta_{10} Size_{i,t-1} + \beta_{11} Institutional_{i,q-1} + \beta_{12} Analyst_{i,q-1} \\ &+ \beta_{13} Buy_{i,q-1} + \beta_{14} Advertising_{i,n-1} + \beta_{15} Earnings_{i,t-1} \\ &+ \beta_{16} Jump_{i,t-1} + \beta_{17} Volume_{i,t-1} + \varepsilon_{i,t,q,n} \end{aligned} \quad (1)$$

where *Investors* is the weekly retail investor base measured with average retail investor base and maximum retail investor base, i denotes stock, t denotes week, q denotes quarter, and n denotes year. *Attention* is the weekly active measures of retail investor attention in week $t-1$, including ASVI and GSVI. We use the lagged weekly measures of retail investor attention to address the potential endogeneity between retail investor attention and retail investor base. *Ret_1W*, *Ret_1M*, and *Ret_3M* are 1-week, 1-month, and 3-month stock returns adjusted with equal-weighted market returns. We use the interaction term between *Attention* and *Ret_1W* to test whether recent stock returns influence retail investors' investment decisions (Chen, 2021; Vozlyublennaia, 2014). *Volatility* is annualized stock return volatility in the prior year. *Illiquidity* is the zero return day ratio in the prior year. *Price* is price level and *BTM* is book-to-market ratio. *Size* is proxied by the natural logarithm of market capitalization. *Institutional* is institutional ownership. *Analyst* and *Buy* are the natural logarithms of the number of analysts following and the number of buy recommendations, respectively. *Advertising* is the natural logarithm of annual advertising expenses in thousands. *Earnings* is the earnings announcement indicator. *Jump* is the stock price jump indicator. *Volume* is the measure for volume shocks. We control for year and stock fixed effects in the estimates. Standard errors are clustered by stock.

If active retail investor attention impacts retail investor base, then an increase in the active measures should predict an increase in retail investor base. We estimate OLS Model (2) to test this hypothesis:

$$\begin{aligned} \Delta Investors_{i,t,q,n} &= \beta_0 + \beta_1 \Delta Attention_{i,t-1} + \beta_2 \Delta Attention_{i,t-1} \\ &* Ret\_1W_{i,t-1} + \beta_3 Ret\_1W_{i,t-1} + \beta_4 Ret\_1M_{i,t-1} + \beta_5 Ret\_3M_{i,t-1} \\ &+ \beta_6 Volatility_{i,n-1} + \beta_7 Illiquidity_{i,n-1} + \beta_8 Price_{i,t-1} + \beta_9 BTM_{i,t-1} \\ &+ \beta_{10} Size_{i,t-1} + \beta_{11} Institutional_{i,q-1} + \beta_{12} Analyst_{i,q-1} \\ &+ \beta_{13} Buy_{i,q-1} + \beta_{14} Advertising_{i,n-1} + \beta_{15} Earnings_{i,t-1} \\ &+ \beta_{16} Jump_{i,t-1} + \beta_{17} Volume_{i,t-1} + \varepsilon_{i,t,q,n} \end{aligned} \quad (2)$$

where $\Delta Investors$ is the weekly change in retail investor base measured with average retail investor base and maximum retail investor base, i denotes stock, t denotes week, q denotes quarter, and n denotes year. Δ Attention is the weekly change in active retail investor attention in week $t-1$, measured with ASVI and GSVI. Other variables are defined in OLS Model (1). We control for year and stock fixed effects in the estimates. Standard errors are clustered by stock. A summary of variables employed in OLS Models (1) and (2) is report in Appendix.

The attention literature conjectures that retail investor attention impacts retail investor base. There is also the view that retail investor base might drive measures of retail investor attention (e.g., Google search frequency). To explore the lead–lag effect between retail investor attention and retail investor base, we follow Hou (2007) and jointly estimate the following vector autoregression (VAR) models:

$$Investors_{i,t} = \beta_0 + \sum_{k=1}^{K} \alpha_k Investors_{i,t-k} + \sum_{k=1}^{K} \beta_k Attention_{i,t-k} + \varepsilon_{i,t} \quad (3)$$

$$Attention_{i,t} = \beta_0 + \sum_{k=1}^{K} \gamma_k Investors_{i,t-k} + \sum_{k=1}^{K} \delta_k Attention_{i,t-k} + \varepsilon_{i,t} \quad (4)$$

where k is the number of lags. Other variables are defined in OLS Model (1). We control for year and stock fixed effects in the estimates. If the impact of retail investor attention on retail investor base is driven by the autocorrelation of retail investor base and the simultaneous interaction between retail investor base and retail investor attention, the coefficients on lags of retail investor attention ($\sum_{k=1}^{K} \beta_k$) should be jointly indistinguishable

[9] We find similar results in unreported tests where raw stock returns and stock returns adjusted with value-weighted market returns are employed, alternatively.

[10] We set the missing values of annual advertising expenses to zero (Albuquerque et al., 2020; Chen et al., 2021).

[11] In unreported tests, we find similar results when volume shocks are measured with the ratio between mean daily trading volume in the week where attention is measured and that in the prior two or four weeks. We also find similar results when stock price jumps in the prior two or four week are controlled in unreported tests.

from zero, when lags of retail investor base are controlled in Model (3).

Similarly, to test the lead–lag effect between the change in retail investor attention and the change in retail investor base, we jointly estimate the following VAR models:

$$\Delta Investors_{i,t} = \beta_0 + \sum_{k=1}^{K} \alpha_k \Delta Investors_{i,t-k} + \sum_{k=1}^{K} \beta_k \Delta Attention_{i,t-k} + \varepsilon_{i,t} \quad (5)$$

$$\Delta Attention_{i,t} = \beta_0 + \sum_{k=1}^{K} \gamma_k \Delta Investors_{i,t-k} + \sum_{k=1}^{K} \delta_k \Delta Attention_{i,t-k} + \varepsilon_{i,t} \quad (6)$$

where all variables are defined in the previous models. We control for year and stock fixed effects in the estimates.

The second critical assumption in the retail investor attention literature is that retail investor base impacts demand for stocks and stock returns. To test this assumption, we examine the impacts of retail investor base on subsequent trading volume, a proxy for stock demand, and stock returns using OLS Model (7):

$$\begin{aligned} KV_{i,t,q,n} &= \beta_0 + \beta_1 Investors_{i,t-1} + \beta_2 Investors_{i,t-1} * Ret\_1W_{i,t-1} \\ &+ \beta_3 Ret\_1W_{i,t-1} + \beta_4 Ret\_1M_{i,t-1} + \beta_5 Ret\_3M_{i,t-1} \\ &+ \beta_6 Volatility_{i,n-1} + \beta_7 Illiquidity_{i,n-1} + \beta_8 Price_{i,t-1} \\ &+ \beta_9 BTM_{i,t-1} + \beta_{10} Size_{i,t-1} + \beta_{11} Institutional_{i,q-1} \\ &+ \beta_{12} Analyst_{i,q-1} + \beta_{13} Buy_{i,q-1} + \beta_{14} Advertising_{i,n-1} \\ &+ \beta_{15} Earnings_{i,t-1} + \beta_{16} Jump_{i,t-1} + \beta_{17} Volume_{i,t-1} + \varepsilon_{i,t,q,n} \end{aligned} \quad (7)$$

where *KV* is our key variables of interest for stock *i*, including mean daily trading volume in millions and stock returns adjusted with equal-weighted market returns in the following 1-week, 2-week, 3-week, and 4-week windows. Other variables are defined in the previous models. We control for year and stock fixed effects in the estimates. Standard errors are clustered by stock.

Ultimately, to explore the impact of the change in investor base on subsequent trading volume and stock returns, we estimate OLS Model (8):

$$\begin{aligned} KV_{i,t,q,n} &= \beta_0 + \beta_1 \Delta Investors_{i,t-1} + \beta_2 \Delta Investors_{i,t-1} \\ &* Ret\_1W_{i,t-1} + \beta_3 Ret\_1W_{i,t-1} + \beta_4 Ret\_1M_{i,t-1} + \beta_5 Ret\_3M_{i,t-1} \\ &+ \beta_6 Volatility_{i,n-1} + \beta_7 Illiquidity_{i,n-1} + \beta_8 Price_{i,t-1} + \beta_9 BTM_{i,t-1} \\ &+ \beta_{10} Size_{i,t-1} + \beta_{11} Institutional_{i,q-1} + \beta_{12} Analyst_{i,q-1} \\ &+ \beta_{13} Buy_{i,q-1} + \beta_{14} Advertising_{i,n-1} + \beta_{15} Earnings_{i,t-1} \\ &+ \beta_{16} Jump_{i,t-1} + \beta_{17} Volume_{i,t-1} + \varepsilon_{i,t,q,n} \end{aligned} \quad (8)$$

where all variables are defined in the previous models. We control for year and stock fixed effects in the estimates. Standard errors are clustered by stock.

4. Empirical results

4.1. Retail investor attention and retail investor base

In this section, we first investigate the impact of active retail investor attention on the size of retail investor base in each of the following four weeks estimating OLS Model (1). If active retail investor attention is a proxy for stock visibility, it should have a positive impact on the size of retail investor base. The pairwise correlation coefficients for the right-hand-side variables reported in Table 2 do not indicate any significant multicollinearity issue. The results for OLS Model (1) are reported in Panel A of Table 3.

Panel A of Table 3 shows a positive relationship between ASVI and average retail investor base in each of the following four weeks, with and without the interaction term, significant at the 1% level in all estimates. These results indicate that stocks with higher levels of active retail investor attention are associated with a larger retail investor base, and the economic significance of this effect appears to decrease only slightly over the following four weeks. This evidence supports the first assumption in the active investor attention literature that active attention positively affects retail investor base. With a higher level of active retail investor attention, winner stocks with a positive 1-week return are associated with an even larger retail investor base, suggesting winner-chasing behavior of retail investors (Ashour et al., 2023; Bailey et al., 2011; Haruvy et al., 2007). This is indicated by the positive relationship between the interaction term and average retail investor base reported in Panel A of Table 3. In addition, stock return in the prior week is positively related to average retail investor base in Columns (1), (3), and (7) of Panel A. However, this result is driven by stocks with a higher level of retail investor attention. When the interaction term is included in Columns (2), (4), (6), and (8), stock return in the prior week is negatively related to average retail investor base. This might be a result of the disposition effect where existing retail investors sell winner stocks too soon (Odean, 1998), leading to a smaller retail investor base. Ultimately, annualized stock return volatility, lower price, information shocks, volume shocks, the number of analysts following a stock, and advertising expenses are positively related to average retail investor base, consistent with the literature (Gervais et al., 2001; Grullon et al., 2004; Lou, 2014; Kumar and Lee, 2006). In general, larger stocks and stocks with higher institutional ownerships tend to have a smaller retail investor base (Kumar and Lee, 2006).

To test whether the increase in active retail investor attention leads to increases in retail investor base, we estimate OLS Model (2) and report results in Panel B of Table 3. The results show a positive relationship between the increase in ASVI and the increase in average retail investor base, with and without the interaction term, significant at the 1% level in all estimates. A one-standard-deviation increase in ASVI increases average retail investor base by approximately 0.14% in each of the subsequent four weeks. This indicates that increases in active attention lead to a larger retail investor base, consistent with the first assumption in the attention literature. The interaction term between the increase in ASVI and 1-week stock return is negatively related to the changes in average retail investor base in weeks 2, 3, and 4, suggesting that loser stocks, with increased active attention, tend to attract more retail investors. This effect could be driven by attracted retail investors "buying the dip" (Dreman et al., 2001; Shohfi and Simaan, 2021) because we find in unreported tests that loser stocks continue to have negative returns in weeks 2 and 3. In addition, stocks with a lower price and smaller stocks tend to experience more significant increases in retail investor base, likely because retail investors tend to concentrate on these stocks (Kumar and Lee, 2006). With our relative measure of retail investor base, stocks with higher institutional ownership tend to experience more significant increases in retail investor base. This might be because stocks with higher institutional ownerships tend to have a smaller retail investor base and a given increase in the number of retail investors will be a relatively more significant increase for these stocks. Buy recommendations, earnings announcements, information shocks, and volume shocks attracts more retail investors.

Moreover, the positive relationship between ASVI and retail investor base documented in this section suggests a positive

Table 2
Pairwise correlation.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1) ASVI	1.0000																		
(2) ΔASVI	**0.4656*****	1.0000																	
(3) Investors	**0.1895*****	0.0017	1.0000																
(4) ΔInvestors	**0.0595*****	**0.0465*****	**0.1394*****	1.0000															
(5) Ret_1W	0.0013	**-0.0100*****	**0.0913*****	**0.0415*****	1.0000														
(6) Ret_1M	**0.0102*****	0.0004	**0.1596******	**0.0356*****	**0.5112*****	1.0000													
(7) Ret_3M	**0.0221*****	0.0018	**0.2120*****	**-0.0258*****	**0.1919*****	**0.4913*****	1.0000												
(8) Volatility	**-0.0194*****	-0.0001	**0.1023*****	**0.0325*****	**0.0405*****	**0.0785*****	**0.1094*****	1.0000											
(9) Illiquidity	**-0.0419*****	-0.0003	**-0.0734*****	**-0.0214*****	**-0.0131*****	**-0.0220*****	**-0.0285*****	**0.1115*****	1.0000										
(10) Price	**-0.0675*****	0.0029	**-0.0084*****	**0.0362*****	**0.0144*****	**0.0161*****	**-0.0135*****	**0.0545*****	**0.0047****	1.0000									
(11) BTM	**-0.0313*****	0.0002	**0.0214*****	**0.0122*****	**0.0140*****	**0.0186*****	0.0009	**-0.0036***	-0.0012	**0.0583*****	1.0000								
(12) Size	**0.3443*****	-0.0014	**0.0556*****	**-0.0077*****	**-0.0085*****	**-0.0068*****	**0.0161*****	**-0.0856*****	**-0.0524*****	**-0.4703*****	**-0.0974*****	1.0000							
(13) Institutional	**0.1699*****	-0.0001	**0.0434*****	**-0.0086*****	-0.0017	**-0.0042****	**-0.0087*****	**-0.0634*****	**-0.0315*****	**-0.3684*****	**-0.0582*****	**0.5533*****	1.0000						
(14) Analyst	**0.3379*****	0.0000	**0.0522*****	0.0021	-0.0005	-0.0008	0.0003	**-0.0322*****	**-0.0415*****	**-0.2870*****	**-0.0709*****	**0.7639*****	**0.5493*****	1.0000					
(15) Buy	**0.2719*****	0.0005	**0.0083*****	**0.0098*****	0.0003	-0.0003	-0.0013	-0.0027	**-0.0269*****	**-0.2312*****	**-0.0899*****	**0.6420*****	**0.4443*****	**0.8520*****	1.0000				
(16) Advertising	**0.1686*****	-0.0003	**0.0281*****	**0.0052*****	0.0011	0.0020	0.0023	**-0.0395*****	**-0.0313*****	**-0.1503*****	**-0.0254*****	**0.3728*****	**0.1995*****	**0.3125*****	**0.2199*****	1.0000			
(17) Earnings	**-0.0043****	**0.0057*****	**-0.0528*****	**0.0372*****	**-0.0055*****	**-0.0107*****	0.0018	**-0.0234*****	**0.0252*****	**-0.0309*****	**-0.0051*****	**0.0202*****	**0.0448*****	**0.0125*****	**0.0122*****	**0.0090*****	1.0000		
(18) Jump	**0.1086*****	**0.0036****	**0.0738*****	**0.0899*****	**-0.0471*****	**-0.0703*****	**-0.0542*****	**-0.0079*****	**-0.0237*****	**-0.0973*****	**-0.0422*****	**0.2781*****	**0.1346*****	**0.1703*****	**0.1756*****	**0.1215*****	**0.0654*****	1.0000	
(19) Volume	**-0.0236*****	**-0.0037****	**-0.0038****	**0.1073*****	-0.0024	**0.0057*****	**0.0087*****	**0.0042****	-0.0016	**0.0202*****	0.0019	**-0.0386*****	**-0.0189*****	**-0.0201*****	**-0.0168*****	**-0.0117*****	**0.0046****	**0.0036****	1.0000

This table reports pairwise correlation coefficients among the right-hand-side variables of OLS Model (1). *Investors* is average retail investor base. *ΔInvestors* is the weekly change in *Investors*. Other variables are defined in Table 1. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

Table 3
Active retail investor attention and retail investor base.

	Panel A — Average retail investor base							
	1st Week		2nd Week		3rd Week		4th Week	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ASVI	**0.0268*****	**0.0266*****	**0.0261*****	**0.0260*****	**0.0253*****	**0.0252*****	**0.0248*****	**0.0246*****
	29.91	*29.82*	*29.04*	*29.00*	*28.08*	*28.05*	*27.68*	*27.65*
Interaction		**0.1124*****		**0.0908*****		**0.0813*****		**0.0749*****
		15.01		*12.52*		*11.09*		*10.19*
Ret_1 W	**0.0415*****	**−0.2112*****	**−0.0247*****	**−0.2288*****	0.0032	**−0.1795*****	**0.0188*****	**−0.1497*****
	12.17	*−12.04*	*−7.40*	*−13.36*	*0.88*	*−10.38*	*5.17*	*−8.64*
Ret_1M	**0.1650*****	**0.1656*****	**0.2044*****	**0.2048*****	**0.2069*****	**0.2073*****	**0.2115*****	**0.2118*****
	29.03	*28.97*	*35.02*	*34.94*	*35.37*	*35.32*	*35.83*	*35.79*
Ret_3M	**0.2734*****	**0.2731*****	**0.2342*****	**0.2340*****	**0.1881*****	**0.1879*****	**0.1466*****	**0.1464*****
	49.15	*49.08*	*42.91*	*42.87*	*35.15*	*35.12*	*27.60*	*27.57*
Volatility	**0.0018*****	**0.0018*****	**0.0018*****	**0.0018*****	**0.0017*****	**0.0017*****	**0.0017*****	**0.0017*****
	12.36	*12.36*	*12.35*	*12.35*	*12.25*	*12.25*	*11.83*	*11.83*
Illiquidity	**−0.1077*****	**−0.1078*****	**−0.1049*****	**−0.1051*****	**−0.1010*****	**−0.1012*****	**−0.0963*****	**−0.0964*****
	−6.05	*−6.07*	*−5.92*	*−5.93*	*−5.70*	*−5.71*	*−5.42*	*−5.43*
Price	**0.0240*****	**0.0241*****	**0.0256*****	**0.0257*****	**0.0270*****	**0.0271*****	**0.0293*****	**0.0294*****
	4.09	*4.11*	*4.34*	*4.36*	*4.58*	*4.59*	*4.94*	*4.96*
BTM	**0.0036***	**0.0036***	**0.0037***	**0.0037***	**0.0038***	**0.0038***	**0.0039***	**0.0039***
	1.88	*1.88*	*1.84*	*1.84*	*1.79*	*1.80*	*1.78*	*1.78*
Size	**−0.0157*****	**−0.0156*****	**−0.0189*****	**−0.0189*****	**−0.0214*****	**−0.0214*****	**−0.0223*****	**−0.0222*****
	−4.08	*−4.08*	*−4.64*	*−4.63*	*−4.96*	*−4.96*	*−5.01*	*−5.00*

(*continued on next page*)

Table 3 (*continued*).

Institutional	**−0.0444*****	**−0.0441*****	**−0.0373*****	**−0.0371*****	**−0.0306****	**−0.0305****	**0.0237***	**−0.0235***
	−3.09	*−3.07*	*−2.60*	*−2.59*	*−2.14*	*−2.13*	*−1.65*	*−1.64*
Analyst	**0.0337*****	**0.0336*****	**0.0300*****	**0.0300*****	**0.0265*****	**0.0265*****	**0.0231****	**0.0231****
	3.49	*3.49*	*3.11*	*3.11*	*2.76*	*2.75*	*2.41*	*2.40*
Buy	**−0.0108***	**−0.0107***	−0.0067	−0.0066	−0.0027	−0.0026	0.0006	0.0007
	−1.72	*−1.72*	*−1.07*	*−1.07*	*−0.43*	*−0.43*	*0.10*	*0.11*
Advertising	**0.0097****	**0.0096****	**0.0094****	**0.0093****	**0.0093****	**0.0092****	**0.0091****	**0.0091****
	2.09	*2.07*	*2.01*	*1.99*	*1.97*	*1.96*	*1.94*	*1.93*
Earnings	**−0.0078*****	**−0.0078*****	**−0.0056*****	**−0.0056*****	**−0.0084*****	**−0.0085*****	**−0.0105*****	**−0.0105*****
	−6.05	*−6.06*	*−4.28*	*−4.29*	*−6.74*	*−6.75*	*−8.57*	*−8.58*
Jump	**0.0179*****	**0.0184*****	**0.0302*****	**0.0305*****	**0.0312*****	**0.0315*****	**0.0335*****	**0.0338*****
	6.88	*7.04*	*11.14*	*11.25*	*11.26*	*11.36*	*11.80*	*11.89*
Volume	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001****	**0.0001****
	3.04	*3.20*	*3.75*	*3.83*	*3.13*	*3.21*	*2.36*	*2.39*
Intercept	**0.5152*****	**0.5158*****	**0.5601*****	**0.5598*****	**0.5947*****	**0.5943*****	**0.6068*****	**0.6064*****
	9.36	*9.37*	*9.63*	*9.63*	*9.74*	*9.75*	*9.68*	*9.68*
F-Test		**467.40*****		**431.20*****		**400.46*****		**388.13*****
N	277,650	277,650	2,75,099	2,75,099	2,72,564	2,72,564	2,70,043	2,70,043
Adj. R^2	0.7223	0.7226	0.7233	0.7235	0.7233	0.7235	0.7248	0.7249

Panel B — Change in Average Retail Investor Base

	1st Week		2nd Week		3rd Week		4th Week	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ ASVI	**0.0012*****	**0.0012*****	**0.0012*****	**0.0013*****	**0.0012*****	**0.0012*****	**0.0012*****	**0.0012*****
	21.34	*21.38*	*18.67*	*18.77*	*16.49*	*16.74*	*16.74*	*17.00*
Interaction		−0.0008		**−0.0073****		**−0.0159*****		**−0.0200*****
		−0.33		*−2.03*		*−3.62*		*−3.74*
Ret_1 W	**0.0097*****	**0.0097*****	−0.0025	−0.0023	0.0017	0.0021	**0.0068****	**0.0072****
	4.41	*4.44*	*−0.94*	*−0.89*	*0.61*	*0.73*	*2.16*	*2.30*
Ret_1M	**0.0232*****	**0.0232*****	**0.0453*****	**0.0454*****	**0.0524*****	**0.0526*****	**0.0528*****	**0.0531*****
	14.94	*14.96*	*15.98*	*16.03*	*12.89*	*12.96*	*10.17*	*10.23*
Ret_3M	**−0.0176*****	**−0.0176*****	**−0.0368*****	**−0.0368*****	**−0.0577*****	**−0.0577*****	**−0.0790*****	**−0.0790*****
	−16.31	*−16.31*	*−17.44*	*−17.44*	*−18.58*	*−18.57*	*−19.67*	*−19.66*
Volatility	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0001*****	**0.0002*****	**0.0002*****
	3.82	*3.82*	*3.74*	*3.73*	*3.53*	*3.53*	*3.50*	*3.49*
Illiquidity	−0.0018	−0.0018	−0.0035	−0.0035	−0.0034	−0.0034	−0.0036	−0.0036
	−1.12	*−1.12*	*−1.16*	*−1.16*	*−0.77*	*−0.77*	*−0.63*	*−0.63*
Price	**0.0026*****	**0.0026*****	**0.0054*****	**0.0054*****	**0.0082*****	**0.0082*****	**0.0106*****	**0.0106*****
	4.77	*4.77*	*5.01*	*5.01*	*5.09*	*5.09*	*5.06*	*5.06*
BTM	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003
	0.66	*0.66*	*0.74*	*0.74*	*0.65*	*0.65*	*0.66*	*0.66*
Size	**−0.0016*****	**−0.0016*****	**−0.0035*****	**−0.0035*****	**−0.0051*****	**−0.0051*****	**−0.0064*****	**−0.0064*****
	−5.81	*−5.80*	*−6.61*	*−6.61*	*−6.56*	*−6.55*	*−6.35*	*−6.35*
Institutional	**0.0038*****	**0.0038*****	**0.0072*****	**0.0072*****	**0.0104*****	**0.0104*****	**0.0124*****	**0.0134*****
	3.82	*3.82*	*3.71*	*3.71*	*3.58*	*3.58*	*3.27*	*3.27*
Analyst	**−0.0046*****	**−0.0046*****	**−0.0092*****	**−0.0092*****	**−0.0139*****	**−0.0139*****	**−0.0183*****	**−0.0183*****
	−7.61	*−7.61*	*−7.85*	*−7.85*	*−7.96*	*−7.96*	*−7.98*	*−7.98*
Buy	**0.0044*****	**0.0044*****	**0.0092*****	**0.0092*****	**0.0140*****	**0.0139*****	**0.0183*****	**0.0183*****
	11.11	*11.11*	*11.60*	*11.60*	*11.71*	*11.71*	*11.62*	*11.61*
Advertising	−0.0001	−0.0001	−0.0002	−0.0002	−0.0001	−0.0001	−0.0001	−0.0001
	−0.53	*−0.53*	*−0.34*	*−0.35*	*−0.16*	*−0.16*	*−0.04*	*−0.04*
Earnings	**0.0098*****	**0.0098*****	**0.0121*****	**0.0121*****	**0.0089*****	**0.0089*****	**0.0067*****	**0.0067*****
	19.04	*19.04*	*15.55*	*15.55*	*12.80*	*12.80*	*9.82*	*9.82*
Jump	**0.0133*****	**0.0133*****	**0.0243*****	**0.0243*****	**0.0255*****	**0.0255*****	**0.0271*****	**0.0271*****
	23.28	*23.28*	*24.44*	*24.44*	*23.46*	*23.48*	*21.54*	*21.56*
Volume	**0.0002*****	**0.0002*****	**0.0003*****	**0.0003*****	**0.0002*****	**0.0002*****	**0.0002*****	**0.0002*****
	4.40	*4.40*	*4.12*	*4.12*	*3.82*	*3.82*	*3.55*	*3.55*
Intercept	**0.0250*****	**0.0249*****	**0.0539*****	**0.0539*****	**0.0803*****	**0.0803*****	**0.1041*****	**0.1041*****
	6.56	*6.56*	*7.39*	*7.39*	*7.46*	*7.46*	*7.38*	*7.38*
F-Test		**228.59*****		**176.36*****		**141.95*****		**145.84*****
N	275,185	275,185	2,72,633	2,72,633	2,70,102	2,70,102	2,67,584	2,67,584
Adj. R^2	0.0635	0.0635	0.0878	0.0878	0.1091	0.1092	0.1365	0.1366

This table reports regression results for OLS Model (1) and OLS Model (2) in Panels A and B, respectively. *Interaction* is the interaction term between *ASVI* and *Ret_1W* in Panel A and between *ΔASVI* and *Ret_1W* in Panel B. Other variables are defined in Table 1. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable (*ASVI* in Panel A and *ΔASVI* in Panel B) and *Interaction* are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

correlation between Google search frequency and the number of Google search users (retail investors). This lends support to recent empirical studies that assume such positive correlation exists (Da et al., 2015; Kostopouls et al., 2020).

4.2. Potential endogeneity

Although previous studies on active retail investor attention assume that active attention impacts retail investor base (Chen, 2017; Chen et al., 2021; Da et al., 2011; Gervais et al., 2001),

Table 4
Endogeneity.

	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(3)	(4)
	Panel A — Simultaneous Interaction Between Attention and Average Retail Investor Base			
Investors $_{t-1}$	**0.9874*****	**0.9641*****	**0.9415*****	**0.9188*****
	1450.08	*658.74*	*434.32*	*325.27*
$ASVI_{t-1}$	**0.0024*****	**0.0029*****	**0.0032*****	**0.0037*****
	24.83	*18.48*	*15.84*	*15.21*
Interaction	−0.0028	−**0.0129*****	−**0.0206*****	−**0.0239*****
	−1.48	*−4.35*	*−5.23*	*−5.17*
Controls	YES	YES	YES	YES
F-Test	**312.33*****	**185.01*****	**143.48*****	**130.86*****
N	275,185	2,72,633	2,70,102	2,67,584
Adj. R^2	0.9875	0.9687	0.9510	0.9343
	Panel B — Simultaneous Interaction Between Changes in Attention and Average Retail Investor Base			
Δ investors $_{t-1}$	**0.2661*****	**0.2304*****	**0.2219*****	**0.2131*****
	45.23	*25.31*	*19.63*	*16.74*
Δ $ASVI_{t-1}$	**0.0008*****	**0.0009*****	**0.0008*****	**0.0009*****
	13.97	*12.65*	*11.38*	*12.13*
Interaction	−0.0012	−**0.0075****	−**0.0163*****	−**0.0203*****
	−0.56	*−2.16*	*−3.74*	*−3.8*
Controls	YES	YES	YES	YES
F-Test	**97.62*****	**80.95*****	**68.87*****	**77.08*****
N	2,72,700	2,70,153	2,67,626	2,65,108
Adj. R^2	0.1292	0.1068	0.1201	0.1440
	Panel C — Average Retail Investor Base Two Weeks Prior			
Investors $_{t-2}$	**0.9609*****	**0.9394*****	**0.9173*****	**0.8946*****
	636.45	*428.69*	*321.58*	*256.27*
$ASVI_{t-1}$	**0.0080*****	**0.0083*****	**0.0084*****	**0.0086*****
	35.18	*31.75*	*28.51*	*26.73*
Interaction	**0.0069****	−0.0042	−**0.0119*****	−**0.0155*****
	2.37	*−1.13*	*−2.67*	*−3.10*
Controls	YES	YES	YES	YES
F-Test	**620.09*****	**512.12*****	**417.58*****	**366.10*****
N	2,72,700	2,70,153	2,67,626	2,65,108
Adj. R^2	0.9689	0.9517	0.9347	0.9187
	Panel D — Change in Average Retail Investor Base Two Weeks Prior			
Δ investors $_{t-2}$	−**0.0343*****	−**0.0433*****	−**0.0518*****	−**0.0627*****
	−7.32	*−5.90*	*−5.69*	*−5.94*
Δ $ASVI_{t-1}$	**0.0012*****	**0.0012*****	**0.0011*****	**0.0012*****
	20.82	*17.94*	*15.69*	*15.82*
Interaction	−0.0010	−**0.0076****	−**0.0164*****	−**0.0205*****
	−0.42	*−2.13*	*−3.73*	*−3.83*
Controls	YES	YES	YES	YES
F-Test	**216.81*****	**161.17*****	**125.52*****	**126.94*****
N	2,70,242	2,67,699	2,65,172	2,62,656
Adj. R^2	0.0642	0.0891	0.1106	0.1383

This table reports regression results for OLS Models (1) and (2). In Panel A and Panel B, lagged (week t-1) average retail investor base and lagged (week t-1) change in average retail investor base are controlled in OLS Models (1) and (2), respectively. In Panel C and Panel D, average retail investor base two weeks prior (week t-2) and the change in average retail investor base two weeks prior (week t-2) are controlled in OLS Models (1) and (2), respectively. *Interaction* is the interaction term between the key variable (*ASVI* in Panel A and Panel C, and *ΔASVI* in Panel B and Panel D) and *Ret_1W*. Other variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

this relationship can be endogenous. For instance, it is possible that retail investor base might drive Google search activities of retail investors. It is also possible that the relationships documented in Table 3 might be driven by the autocorrelation in retail investor base data and the simultaneous interaction between retail investor base and active retail investor attention.[12] These possibilities are worth investigating, even though we use lagged measures of active retail investor attention (in week $t-1$) in the previous section to address the potential endogeneity.

If the effect of lagged active retail investor attention on retail investor base is driven by the simultaneous interaction between active retail investor attention and retail investor base, and the autocorrelation in retail investor base data, this effect should disappear when lagged retail investor base is controlled in the estimates. We therefore estimate OLS Model (1), adding lagged average retail investor base (in week $t-1$) as an independent variable, and report results for variables of interest in Panel A of Table 4. Similarly, we add the lagged change in average retail investor base (in week $t-1$) as an independent variable and estimate OLS Model (2). The results for the variables of interest are reported in Panel B of Table 4.

[12] We find in unreported tests that (the change in) retail investor base is positively related to (the change in) retail investor attention in the same week, significant at the (1%) 1% level.

Table 5
Lead–lag effects.

Dependent variable	Panel A — One-lag Regressions		Panel B — Two-lag Regressions	
	Investors (t-1)	ASVI (t-1)	Investors (t-1: t-2)	ASVI (t-1: t-2)
Investors	**0.9865*****	**0.0024*****	**0.9788*****	**0.0010*****
	6048.53	*1226.91*	*6111.78*	*100.21*
ASVI	**0.5859*****	**0.1202*****	**0.4735*****	**0.1830*****
	2694.14	*3896.68*	*1694.02*	*4957.10*
Dependent variable	Panel C — Three-lag Regressions		Panel D — Four-lag Regressions	
	Investors (t-1: t-3)	ASVI (t-1: t-3)	Investors (t-1: t-4)	ASVI (t-1: t-4)
Investors	**0.9828*****	**0.0015*****	**0.9799*****	**0.0016*****
	5859.03	*247.65*	*5535.81*	*177.59*
ASVI	**0.4087*****	**0.2311*****	**0.3508*****	**0.2707*****
	1200.95	*5664.90*	*843.20*	*6159.01*
Dependent variable	Panel E — One-lag Regressions		Panel F — Two-lag Regressions	
	Δ Investors (t-1)	Δ ASVI (t-1)	Δ Investors (t-1: t-2)	Δ ASVI (t-1: t-2)
Δ Investors	**0.2701*****	**0.0008*****	**0.1879*****	**0.0009*****
	2036.93	*233.28*	*6159.67*	*70.84*
Δ ASVI	**−0.3855*****	**−0.4848*****	**−0.2479*****	**−0.9603*****
	36.17	*8146.25*	*10.34*	*8985.45*
Dependent variable	Panel G — Three-lag Regressions		Panel H — Four-lag Regressions	
	Δ Investors (t-1: t-3)	Δ ASVI (t-1: t-3)	Δ Investors (t-1: t-4)	Δ ASVI (t-1: t-4)
Δ Investors	**0.2184*****	**0.0014*****	**0.2063*****	**0.0018*****
	5623.73	*77.16*	*3798.57*	*72.84*
Δ ASVI	−0.1006	**−1.4296*****	−0.0189	**−1.8862*****
	1.22	*8563.83*	*0.03*	*7795.05*

This table reports the VARs results for the lead–lag effect between retail investor attention and retail investor base in Panels A through D (Model (3) and Model (4)), and the VARs results for the lead–lag effect between the change in retail investor attention and the change in retail investor base in Panels E though H (Model (5) and Model (6)). The VARs are estimated with one through four lags and the sum of the coefficients on lagged variables is reported. In this table, *t* denotes week and other variables are defined in the previous tables. F-statistics (t-statistics) for VARs estimates with multiple lags (one lag) are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

The results reported in Panel A of Table 4 show that, when lagged average retail investor base is controlled in the estimates, ASVI is still positively related to average retail investor base in each of the following four weeks, significant at the 1% level in all estimates. However, the economic significance of this effect is lower than that in Panel A of Table 3. This reduction is not surprising given the simultaneous interaction between active retail investor attention and retail investor base, and the autocorrelation in active retail investor attention data.[13] The results reported in Panel B also show a positive relationship between the change in ASVI and the change in average retail investor base in the following four weeks when lagged change in average retail investor base is controlled in the estimates, significant at the 1% level in all estimates. The economic significance of this effect (around 0.0009) is similar to that reported in Panel B of Table 3 (0.0012).

To test whether the results reported in Table 3 are driven by the impact of lagged retail investor base (week $t-2$) on active retail investor attention (week $t-1$), we estimate OLS Model (1) adding average retail investor base two weeks prior (week $t-2$) as an independent variable. Similarly, we add the change in average retail investor base two weeks prior (week $t-2$) as an independent variable and estimate OLS Model (2). Results for variables of interest are reported in Panel C and Panel D of Table 4.

The results reported in Panel C of Table 4 suggest that those reported in Table 3 are not driven by the impact of lagged retail investor base on active retail investor attention. When the average retail investor base from two weeks prior is controlled in the estimates, ASVI is still positively related to average retail investor base in the following four weeks, significant at the 1% level in all estimates. Panel D of Table 4 shows that when the change in average retail investor base from two weeks prior is controlled in the estimates, the change in ASVI is still positively related to the change in average retail investor base in the following four weeks, significant at the 1% level in all estimates. The results for the interaction terms in Panel C and Panel D of Table 4 indicate the influence of recent stock returns on the effect of active retail investor attention on retail investor base. In total, we find that our results are not driven by the impact of lagged retail investor base on active retail investor attention.

Ultimately, we estimate VAR models (Model (3) and Model (4)) to explore the lead–lag effect between retail investor attention and retail investor base. Results for one through four lags are reported in Panels A through D of Table 5, respectively. Similarly, we estimate VAR models (Model (5) and Model (6)) to test the lead–lag effect between the change in retail investor attention and the change in retail investor base. Results for one through four lags are reported in Panels E through H of Table 5, respectively.

The results reported in Panels A through D of Table 5 show that ASVI positively impacts retail investor base in VAR estimates with one through four lags, significant at the 1% level in all estimates. They also show that lagged retail investor base positively influences ASVI. The results reported in Panels E though H indicate that the increase in ASVI drives the increase in retail investor base in VAR estimates with one through four lags, significant at the 1% level in all estimates.

To conclude, the results reported in Tables 4 and 5 suggest that the positive impact of active retail investor attention on retail investor base is not driven by the endogeneity between retail investor base and active retail investor attention. This evidence provides critical support to previous studies arguing active retail investor attention positively impacts retail investor base.

[13] We find in unreported tests that active retail investor attention is positively related to lagged active retail investor attention, significant at the 1% level.

Table 6
Subsequent trading volumes.

	Panel A — Average Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Investors	**1.6494*****	**1.4418*****	**1.2860*****	**1.1590*****
	7.68	*6.98*	*6.47*	*6.05*
Interaction	0.6430	0.3398	0.6098	0.3656
	0.64	*0.34*	*0.64*	*0.36*
Controls	YES	YES	YES	YES
F-Test	**30.58*****	**25.74*****	**22.22*****	**20.03*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.6333	0.6823	0.7100	0.7279
	Panel B — Change in Average Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Δ Investors	**7.3967*****	**6.0147*****	**5.2317*****	**4.6743*****
	11.48	*10.96*	*10.23*	*9.57*
Interaction	−**21.3725****	−10.0908	−10.0390	−8.3305
	−2.19	*−1.27*	*−1.60*	*−1.39*
Controls	YES	YES	YES	YES
F-Test	**66.26*****	**60.74*****	**53.52*****	**46.61*****
N	275,185	275,185	275,185	275,185
Adj. R^2	0.6329	0.6818	0.7098	0.7277

This table reports regression results of OLS Models (7) and (8) where the dependent variables are mean daily trading volume in millions over the subsequent 1-week, 2-week, 3-week, and 4-week windows. Panel A presents results for OLS Model (7) and Panel B for Model (8). *Interaction* is the interaction term between the key variable (*Investors* in Panel A and *ΔInvestors* in Panel B) and *Ret_1W*. Other variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

4.3. Subsequent trading volumes

In the previous sections, we find evidence that active retail investor attention impacts retail investor base. In this section, we estimate OLS Models (7) and (8) to test the second assumption in the investor attention literature where retail investor base impacts stock demand. Results are reported in Panels A and B of Table 6, respectively.[14] We use the mean daily trading volume over the subsequent 1-week, 2-week, 3-week, and 4-week windows to proxy stock demand. Only results for variables of interest are reported for brevity.

The results reported in Panel A of Table 6 show a positive impact of average retail investor base on subsequent mean daily trading volume over the four windows, all significant at the 1% level. The results reported in Panel B of Table 6 suggest that the increase in average retail investor base leads to higher mean daily trading volume over the four windows, all significant at the 1% level. A one-standard-deviation increase in average retail investor base increases mean daily trading volume by approximately 0.2278, 0.1853, 0.1611, and 0.1440 million over the following 1-week, 2-week, 3-week, and 4-week windows, respectively. In addition, the interaction term between the change in average retail investor base and 1-week stock return has a negative impact on mean daily trading volume in the subsequent week.

The results reported in Table 6 support prior investor attention studies that postulate a positive relationship between investor base and stock demand (Chen, 2021; Kaniel et al., 2012; Moore, 2020), supporting the buying pressure hypothesis.

4.4. Subsequent stock returns

In this section, we estimate OLS Models (7) and (8) to examine the second assumption in the attention literature where retail investor base and the change in retail investor impact subsequent stock returns. We employ stock returns adjusted with equal-weighted market returns over the subsequent 1-week, 2-week, 3-week, and 4-week windows in our analysis. We report the results for average retail investor base and the change in average retail investor base in Panels A and B of Table 7, respectively.[15] Only the results for variables of interest are reported for brevity.

The results reported in both panels of Table 7 show that average retail investor base and the change in average retail investor base are positively related to adjusted stock returns over the subsequent four windows, all significant at the 1% level. Panel B of Table 7 shows that a one-standard-deviation increase in average retail investor base increases adjusted stock returns by approximately 0.20%, 0.35%, 0.53%, and 0.62% over the subsequent four windows, respectively. This is potentially caused by the price pressure due to higher demand (Ben-Rephael et al., 2017; Mayer, 2021), as documented in Table 6. In addition, the two interaction terms in Table 7 are negatively related to subsequent adjusted stock returns, significant at the 1% level in all estimates. This evidence suggests that, with a larger retail investor base, winner stocks tend to have a lower return in the following weeks. This may be an indication of the disposition effect where retail investors tend to sell winner stocks (Odean, 1998; Rau, 2015).

Overall, the results reported in Tables 3–7 show that active retail investor attention impacts retail investor base, which influences stock demand and stock returns. This evidence provides

[14] In unreported tests, we repeat our analysis employing retail investor base and changes in retail investor base predicted by OLS Models (1) and (2), respectively, and find consistent results.

[15] In unreported tests, we find consistent results employing retail investor base and changes in retail investor base predicted by OLS Models (1) and (2), respectively.

Table 7
Subsequent stock returns.

	Panel A — Average Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Investors	**0.0343*****	**0.0653*****	**0.1023*****	**0.1289*****
	48.06	*50.26*	*50.73*	*51.01*
Interaction	**−0.3885*****	**−0.3204*****	**−0.4474*****	**−0.1469*****
	−42.99	*−35.00*	*−31.77*	*−16.43*
Controls	YES	YES	YES	YES
F-Test	**1287.97*****	**1263.74*****	**1297.80*****	**1610.74*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.0384	0.1016	0.1206	0.1466
	Panel B — Change in Average Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Δ Investors	**0.0647*****	**0.1140*****	**0.1708*****	**0.2004*****
	13.10	*21.31*	*22.10*	*25.03*
Interaction	**−2.6872*****	**−1.9040*****	**−2.9035*****	**−1.9884*****
	−13.06	*−12.22*	*−11.40*	*−10.43*
Controls	YES	YES	YES	YES
F-Test	**93.92*****	**235.13*****	**287.69*****	**340.54**
N	275,185	275,185	275,185	275,185
Adj. R^2	0.0272	0.0726	0.0789	0.0956

This table reports regression results of OLS Models (7) and (8) where the dependent variables are stock returns adjusted with equal-weighted market returns over the subsequent 1-week, 2-week, 3-week, and 4-week windows. Panel A presents results for OLS Model (7) and Panel B for OLS Model (8). *Interaction* is the interaction term between the key variable (*Investors* in Panel A and *ΔInvestors* in Panel B) and *Ret_1W*. Other variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

a mechanism through which active retail investor attention impacts stock returns (Chen et al., 2021; Chemmanur and Yan, 2019; Da et al., 2011; Drake et al., 2017; Gervais et al., 2001; Kaniel et al., 2012; Lou, 2014; Mayer, 2021; Swamy and Dharani, 2019).

4.5. The size effect

Previous studies suggest that stock size influences the impacts of retail investor attention (Chemmanur and Yan, 2019; Meshcheryakov and Winters, 2022; Rakowski et al., 2021). To examine the role of stock size in active retail investor attention, we divide our sample into smaller (bottom 30%) and larger (top 30%) size groups and repeat our analyses for the two groups.[16] For brevity, only results for variables of interest are reported in Table 8.[17]

The comparison between the smaller and larger size groups presented in Panel A of Table 8 indicates that ASVI and the interaction term between ASVI and 1-week stock return have economically more significant impacts on average retail investor base for the larger size group in each of the subsequent four weeks. Panel B of Table 8 also shows that the change in ASVI has an economically more significant impact on the change in average retail investor base for the larger size group. In addition, the influence of the interaction term between the change in ASVI and 1-week stock return on the change in average retail investor base, as documented in Panel B of Table 3, is driven by the larger size group.

Panel C of Table 8 shows that average retail investor base and the change in average retail investor base have economically more significant impacts on mean daily trading volume of the larger size group over the subsequent four windows. These results are consistent with Meshcheryakov and Winters (2022) where retail investors trade more actively in larger stocks, and increases in active retail investor attention lead to more significant increases in retail investors' trading of larger stocks. In addition, the interaction term between average retail investor base and 1-week stock return appears to have a positive impact on mean daily trading volume of the larger size group but a negative impact on that of the smaller size group. This means that, with a larger retail investor base, smaller winner stocks tend to experience a decrease in subsequent mean daily trading volume and, larger winner stocks tend to experience an increase in subsequent mean daily trading volume. This might be because the disposition effect dominates the winner-chasing effect for smaller stocks, but has the opposite effect for larger stocks.

As shown in Panel D of Table 8, average retail investor base and the change in average retail investor base have economically more significant impacts on adjusted stock returns of the larger stock size group over the subsequent four windows. For instance, a one-standard-deviation increase in average retail investor base increases adjusted stock returns by 0.12%, 0.22%, 0.36%, and 0.47% for the smaller size group over the subsequent 1-week, 2-week, 3-week, and 4-week windows, and by 0.38%, 0.73%, 1.12%, and 1.24% for the larger size group over the four windows. Moreover, the interaction terms in Panel D of Table 8 also have economically more significant impacts on subsequent adjusted stock returns of the larger size group.

5. Robustness tests

5.1. Alternative measure of retail investor base

In the previous sections, weekly retail investor base is measured with average retail investor base. In this section, we use maximum retail investor base to measure weekly retail investor base in our analyses and report the results for OLS Models (1), (2), (7), and (8) in Table 9. Only results for variables of interest are reported for brevity.

[16] We find consistent results in unreported tests when alternative divisors are used, e.g., bottom and top 40%.

[17] We find consistent results in unreported tests where maximum retail investor base and GSVI are employed to measure retail investor base and retail investor attention, respectively.

Table 8
The size effect.

	Panel A — Average Retail Investor Base							
	Smaller Group				Larger Group			
	1st Week	2nd Week	3rd Week	4th Week	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(2)	(4)	(5)	(6)	(7)	(8)
ASVI	**0.0227*****	**0.0218*****	**0.0208*****	**0.0201*****	**0.0362*****	**0.0365*****	**0.0360*****	**0.0358*****
	18.24	*17.77*	*17.17*	*16.87*	*14.86*	*14.68*	*14.25*	*14.15*
Interaction	**0.0599*****	**0.0474*****	**0.0444*****	**0.0388*****	**0.1693*****	**0.1361*****	**0.1291*****	**0.1228*****
	4.65	*3.79*	*3.58*	*3.13*	*8.41*	*6.97*	*14.25*	*6.03*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**168.67*****	**159.43*****	**148.83*****	**143.82*****	**116.65*****	**110.41*****	**103.76*****	**101.76*****
N	83,295	83,295	83,295	83,295	83,295	83,295	83,295	83,295
Adj. R^2	0.7518	0.7540	0.7545	0.7555	0.7495	0.7500	0.7498	0.7516
	Panel B — Change in Average Retail Investor Base							
	Smaller Group				Larger Group			
	1st Week	2nd Week	3rd Week	4th Week	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(2)	(4)	(5)	(6)	(7)	(8)
Δ ASVI	**0.0012*****	**0.0012*****	**0.0011*****	**0.0011*****	**0.0017*****	**0.0017*****	**0.0018*****	**0.0020*****
	12.82	*11.13*	*10.03*	*9.38*	*12.31*	*11.15*	*10.09*	*10.46*
Interaction	0.0049	0.0025	−0.0062	−0.0112	**−0.0110***	**−0.0298*****	**−0.0397*****	**−0.0493*****
	1.11	*0.43*	*−0.90*	*−1.39*	*−1.83*	*−3.22*	*−3.37*	*−3.72*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**83.19*****	**63.01**	**50.30*****	**43.98*****	**75.86*****	**62.76*****	**52.53*****	**55.77*****
N	82,556	82,556	82,556	82,556	82,556	82,556	82,556	82,556
Adj. R^2	0.0752	0.0995	0.1197	0.1487	0.1209	0.1608	0.1842	0.2177
	Panel C — Mean Daily Trading Volume in Millions							
	Smaller Group				Larger Group			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)	(5)	(6)	(7)	(8)
				Average Retail Investor Base				
Investors	**1.0785*****	**0.8648*****	**0.7206*****	**0.6210*****	**2.5627*****	**2.2178*****	**1.9536*****	**1.7126*****
	5.43	*4.68*	*4.07*	*3.64*	*4.71*	*4.19*	*3.82*	*3.49*
Interaction	−1.2164	−2.0599	**−2.4404***	**−3.5699****	**6.5242*****	**6.1491****	**6.7911*****	**7.4890*****
	−0.74	*−1.27*	*−1.74*	*−2.35*	*2.47*	*2.37*	*2.55*	*2.68*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**15.30*****	**11.00*****	**8.65*****	**8.26*****	**11.20*****	**9.01*****	**7.67*****	**6.66*****
N	83,295	83,295	83,295	83,295	83,295	83,295	83,295	83,295
Adj. R^2	0.1797	0.2365	0.2777	0.3071	0.7405	0.7697	0.7855	0.7963
				Change in Average Retail Investor Base				
Δ Investors	**9.0101*****	**6.9836*****	**5.8967*****	**5.0614*****	**11.9953*****	**10.2353*****	**9.0982*****	**8.4691*****
	6.90	*6.80*	*6.17*	*5.49*	*6.06*	*5.46*	*5.04*	*4.82*
Interaction	−2.6529	−5.1821	−7.0598	−1.2981	−25.2939	7.5359	2.0330	1.6104
	−0.11	*−0.24*	*−0.44*	*−0.09*	*−1.46*	*0.55*	*0.17*	*0.14*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**23.82*****	**23.56*****	**19.93*****	**15.61*****	**20.11*****	**15.02*****	**12.83*****	**11.65*****
N	82,556	82,556	82,556	82,556	82,556	82,556	82,556	82,556
Adj. R^2	0.1896	0.2462	0.2878	0.3168	0.7392	0.7687	0.7849	0.7959
	Panel D — Adjusted Stock Returns							
	Smaller Group				Larger Group			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)	(5)	(6)	(7)	(8)
				Average Retail Investor Base				
Investors	**0.0315*****	**0.0605*****	**0.0943*****	**0.1193*****	**0.0412*****	**0.0777*****	**0.1212*****	**0.1516*****
	23.17	*24.71*	*25.04*	*25.13*	*29.96*	*31.65*	*31.83*	*31.93*
Interaction	**−0.3221*****	**−0.2772*****	**−0.3649*****	**−0.1306*****	**−0.5182*****	**−0.4404*****	**−0.6170*****	**−0.2173*****
	−22.81	*−17.02*	*−14.90*	*−8.00*	*−27.76*	*−22.94*	*−20.95*	*−10.85*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**332.84*****	**307.43*****	**314.76*****	**346.59*****	**517.48*****	**508.53*****	**507.22*****	**569.23*****
N	83,295	83,295	83,295	83,295	83,295	83,295	83,295	83,295
Adj. R^2	0.0379	0.1009	0.1235	0.1543	0.0579	0.1167	0.1406	0.1660

(*continued on next page*)

The results reported in Panel A of Table 9 show that ASVI is positively related to maximum retail investor base in each of the subsequent four weeks, significant at the 1% level in all estimates. Panel B of Table 9 shows that the change in ASVI is positively related to the change in maximum retail investor base in the following week, significant at the 1% level. These results, consistent with those reported in Table 3, support the first assumption in the investor attention literature that retail investor attention is positively related to retail investor base.

The results presented in Panel C of Table 9 are consistent with those reported in Table 6. Maximum retail investor base and the change in maximum retail investor base are positively related to mean daily trading volume over the subsequent 1-week, 2-week, 3-week, and 4-week windows, significant at the 1% level

Table 8 (*continued*).

	Change in Average Retail Investor Base							
Δ Investors	**0.0399*****	**0.0719*****	**0.1159*****	**0.1542*****	**0.1220*****	**0.2384*****	**0.3637*****	**0.4042*****
	4.61	*8.20*	*9.52*	*12.58*	*12.07*	*17.41*	*17.56*	*17.37*
Interaction	**−1.7700*****	**−1.1848*****	**−1.4623*****	**−0.7360****	**−4.8015*****	**−3.2376*****	**−5.5798*****	**−4.2396*****
	−4.55	*−4.31*	*−3.16*	*−2.17*	*−15.17*	*−12.08*	*−12.69*	*−12.06*
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	**11.25*****	**36.50*****	**76.17*****	**105.75*****	**127.96*****	**155.84*****	**158.21*****	**156.74*****
N	82,556	82,556	82,556	82,556	82,556	82,556	82,556	82,556
Adj. R^2	0.0257	0.0759	0.0880	0.1134	0.0431	0.0872	0.1012	0.1127

This table reports regression results of OLS Models (1), (2), (7) and (8) where the sample is divided into smaller (bottom 30%) and larger (top 30%) size groups. Results for OLS Models (1) and (2) are reported in Panels A and B, respectively. OLS Models (7) and (8) results for mean daily trading volume and adjusted stock returns are reported in Panels C and D, respectively. All variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

Table 9
Alternative measure of retail investor base.

	Panel A — Maximum Retail Investor Base			
	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(3)	(4)
ASVI	**0.0284*****	**0.0269*****	**0.0258*****	**0.0250*****
	30.25	*28.73*	*27.65*	*27.06*
Interaction	**0.1113*****	**0.0913*****	**0.0783*****	**0.0754*****
	13.22	*9.08*	*8.54*	*7.53*
Controls	YES	YES	YES	YES
F-Test	**483.87*****	**435.47*****	**394.45*****	**383.31*****
N	277,650	2,75,099	2,72,564	2,70,043
Adj. R^2	0.6695	0.6696	0.6699	0.6713
	Panel B — Change in Maximum Retail Investor Base			
	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(3)	(4)
Δ ASVI	**0.0004*****	0.0001	−0.0001	0.0001
	4.18	*1.29*	*−0.05*	*0.40*
Interaction	**−0.0063****	**−0.0121*****	**−0.0227*****	**−0.0263*****
	−1.96	*−2.66*	*−4.11*	*−3.97*
Controls	YES	YES	YES	YES
F-Test	**11.13*****	**4.25*****	**8.46*****	**7.89*****
N	275,185	2,72,633	2,70,102	2,67,584
Adj. R^2	0.0422	0.0436	0.0627	0.0877
	Panel C — Mean Daily Trading Volume in Millions			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(3)	(4)
Investors	**1.3427*****	**1.1566*****	**1.0251*****	**0.9208*****
	4.17	*4.03*	*3.91*	*3.80*
Interaction	−0.4153	−0.4481	−0.1826	−0.3111
	−0.44	*−0.50*	*−0.21*	*−0.36*
Controls	YES	YES	YES	YES
F-Test	**13.46*****	**12.16*****	**10.96*****	**10.10*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.6331	0.6820	0.7098	0.7277
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(5)	(6)	(7)	(8)
Δ Investors	**3.0471*****	**2.3480*****	**2.0107*****	**1.7633*****
	3.74	*3.69*	*3.70*	*3.65*
Interaction	**−18.2981*****	**−11.5845****	**−12.0933*****	**−9.9771****
	−2.57	*−2.00*	*−2.56*	*−2.22*
Controls	YES	YES	YES	YES
F-Test	**10.88*****	**9.52*****	**10.67*****	**9.61*****
N	275,185	275,185	275,185	275,185
Adj. R^2	0.6318	0.6809	0.7091	0.7271

(*continued on next page*)

in all estimates. In addition, the interaction term between the change in maximum retail investor base and 1-week stock return is negatively related to mean daily trading volume over the four windows, significant at the 1% level in all estimates.

The results reported in Panel D of Table 9 are consistent with those reported in Table 7. Maximum retail investor base and the change in maximum retail investor base have a positive impact on adjusted stock returns over the subsequent 1-week, 2-week, 3-week, and 4-week windows, significant at the 1% level in all estimates. The two interaction terms in Panel D of Table 9 have a negative impact on subsequent adjusted stock returns. These results further support the second assumption in the investor attention literature that retail investor base positively impacts stock demand and returns.

Table 9 (*continued*).

	Panel D — Adjusted Stock Returns			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(3)	(4)
Investors	**0.0269*****	**0.0504*****	**0.0789*****	**0.0987*****
	6.13	*5.64*	*5.60*	*5.41*
Interaction	**−0.3567*****	**−0.2983*****	**−0.4277*****	**−0.1822*****
	−11.81	*−18.45*	*−19.48*	*−7.76*
Controls	YES	YES	YES	YES
F-Test	**218.77*****	**257.30*****	**242.58*****	**328.51*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.0351	0.0946	0.0110	0.1346
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(5)	(6)	(7)	(8)
Δ Investors	**0.0234*****	**0.0385*****	**0.0591*****	**0.0676*****
	6.33	*4.85*	*4.53*	*4.51*
Interaction	**−1.3479*****	**−0.9060*****	**−1.4872*****	**−1.1838*****
	−11.20	*−10.31*	*−9.68*	*−9.71*
Controls	YES	YES	YES	YES
F-Test	**65.00*****	**55.24*****	**47.93*****	**53.43*****
N	275,185	275,185	275,185	275,185
Adj. R^2	0.0203	0.0681	0.0734	0.0912

This table reports regression results of OLS Models (1), (2), (7), and (8) where the dependent variables are maximum retail investor base in Panels A, the change in maximum retail investor base in Panel B, mean daily trading volumes over the subsequent 1-week, 2-week, 3-week, and 4-week windows in Panel C, and stock returns adjusted with equal-weighted market returns over the subsequent 1-week, 2-week, 3-week, and 4-week windows in Panel D. Weekly retail investor base is measured with maximum retail investor base. All variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

Table 10
Alternative measure of active retail investor attention.

	Panel A — Average Retail Investor Base			
	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(2)	(4)
GSVI	**0.0157*****	**0.0153*****	**0.0148*****	**0.0145*****
	26.97	*26.30*	*25.46*	*25.20*
Interaction	**0.0695*****	**0.0558*****	**0.0485*****	**0.0447*****
	12.58	*10.43*	*8.87*	*8.10*
Controls	YES	YES	YES	YES
F-Test	**383.04*****	**354.59*****	**329.42*****	**321.74*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.7193	0.7204	0.7205	0.7221
	Panel B — Change in Average Retail Investor Base			
	1st Week	2nd Week	3rd Week	4th Week
	(1)	(2)	(2)	(4)
Δ GSVI	**0.0007*****	**0.0007*****	**0.0007*****	**0.0007*****
	18.45	*15.31*	*13.28*	*14.60*
Interaction	0.0004	−0.0015	**−0.0061***	**−0.0076****
	0.23	*−0.56*	*−1.89*	*−1.97*
Controls	YES	YES	YES	YES
F-Test	**170.79*****	**117.19*****	**88.52*****	**106.80*****
N	275,185	275,185	275,185	275,185
Adj. R^2	0.0626	0.0875	0.1089	0.1364

This table reports regression results of OLS Models (1) and (2) in Panels A and B, respectively. In this table, active retail investor attention is measured with GSVI and retail investor base is measured with average retail investor base. All variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

5.2. Alternative definition of retail investor attention

In the previous sections, active retail investor attention is measured with ASVI. We measure active retail investor attention with GSVI and estimate OLS Models (1) and (2) in this section. Results for OLS Models (1) and (2) are reported in Panels A and B of Table 10, respectively. For brevity, only results for variables of interest are reported.

The results reported in Panel A of Table 10 are consistent with those reported in Panel A of Table 3. GSVI is positively related to average retail investor base in each of the following four weeks. The interaction term between GSVI and 1-week stock return is also positively related to average retail investor base in each of the following four weeks. These relationships are significant at the 1% level in all estimates. The results reported in Panel B are consistent with those reported in Panel B of Table 3. The change in GSVI is positively related to changes in average retail investor base in the following four weeks, significant at the 1% level in all estimates. A one-standard-deviation increase in GSVI is associated with an increase of approximately 0.11% in average retail investor base in each of the following four weeks. We

Table 11
Subsequent annualized stock return volatilities.

	Panel A — Weekly Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Investors	**1.3888*****	0.0947	**−1.2828*****	**−2.4275*****
	7.07	*0.42*	*−5.23*	*−9.15*
Interaction	**37.5517*****	**34.6799*****	**33.4260*****	**26.1075*****
	8.93	*8.80*	*7.87*	*7.12*
Controls	YES	YES	YES	YES
F-Test	**67.59*****	**39.14*****	**42.15*****	**60.61*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.2177	0.2811	0.3148	0.3431
	Panel B — Change in Weekly Retail Investor Base			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
Δ Investors	**12.3405*****	**9.4469*****	**6.3358*****	**4.9218*****
	16.28	*13.48*	*9.88*	*7.95*
Interaction	**36.8265****	**87.0809*****	**104.2621*****	**76.4731*****
	1.94	*5.78*	*7.49*	*6.20*
Controls	YES	YES	YES	YES
F-Test	**145.51*****	**125.76*****	**95.21*****	**60.60*****
N	275,185	275,185	275,185	275,185
Adj. R^2	0.2181	0.2813	0.3142	0.3414

This table reports regression results of OLS Models (7) and (8) where the dependent variables are annualized stock return volatilities over the subsequent 1-week, 2-week, 3-week, and 4-week windows. Panel A presents results for OLS Model (7) and Panel B for OLS Model (8). All variables are defined in the previous tables. Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

find consistent results in unreported tests where investor base is measured with maximum retail investor base.

5.3. Subsequent stock return volatilities

Previous studies find retail investor attention impacts stock return volatilities (Andrei and Hasler, 2015; Aouadi et al., 2013; Herwartz and Xu, 2022) but the results are mixed. In this section, we estimate OLS Models (7) and (8) to test whether retail investor base and the change in retail investor base influence subsequent annualized stock returns volatilities over the subsequent 1-week, 2-week, 3-week, and 4-week windows. We report the results for average retail investor base and the change in average retail investor base in Panels A and B of Table 11, respectively. Only results for variables of interest are reported for brevity. We find consistent results in unreported tests where retail investor base is measured with maximum retail investor base.

The results reported in the two panels of Table 11 show that both interaction terms are positively related to annualized stock return volatilities over the four windows, significant at the 1% level in all estimates. This indicates that average retail investor base and the change in average retail investor base have a positive impact on subsequent annualized stock return volatilities when 1-week stock returns are positive.

The change in average retail investor base also has a positive impact on subsequent annualized stock return volatilities, significant at the 1% level in all estimates in Panel B. A one-standard-deviation increase in average retail investor base increases annualized stock return volatilities by approximately 0.3801, 0.2910, 0.1951, and 0.1516 over the 1-week, 2-week, 3-week, and 4-week windows, respectively. These results are consistent with Jankensgård and Vilhelmsson (2018) where retail ownership is positively related to stock return volatilities in the Swedish market. In contrast, average retail investor base in Panel A has a positive impact on annualized stock return volatilities over the 1-week window, but a negative impact over the 3-week and 4-week windows.

5.4. The impacts of active retail investor attention

In this study, we aim to provide empirical support for the critical assumptions in prior studies documenting positive impacts of active retail investor attention on stock demand and stock returns. We cannot apply our main conclusions to those prior studies unless we find results within our framework that are consistent with those in the prior studies. Therefore, we estimate OLS Models (1) and (2) to test the impacts of active retail investor attention on mean daily trading volume, stock returns adjusted with equal-weighted market returns, and annualized stock returns volatilities over the subsequent 1-week, 2-week, 3-week, and 4-week windows (dependent variables). Results for ASVI and the change in ASVI are reported in Panel A and Panel B of Table 12, respectively. Only results for key variables of interest are reported.

The results reported in Panel A of Table 12 show that ASVI has a positive impact on mean daily trading volume, adjusted stock returns, and annualized stock return volatilities over the subsequent four windows, significant at the 1% level in most estimates. These results are consistent with the findings in the previous studies (Ding and Hou, 2015). The interaction term between ASVI and 1-week stock return is negatively related to mean daily trading volume, adjusted stock returns, and annualized stock return volatilities over the subsequent four windows, significant at the 1% level in most estimates.

Consistent with the findings in the previous studies (Kim et al., 2019; Takeda and Wakao, 2014), the results reported in Panel B of Table 12 show a positive impact of the change in ASVI on mean daily trading volume and annualized stock return volatilities over the subsequent four windows, significant at the 1% level in all estimates. In contrast to Da et al. (2011), the change in ASVI is negatively related to adjusted stock returns over the subsequent four windows, significant at the 1% level in all estimates. However, this result is consistent with more recent studies (Bijl et al., 2016; Chen, 2017; Chen et al., 2020) where increases in active retail investor attention indicate investor overreaction (Heyman et al., 2019) and therefore, predicts decreases in stock prices.

Table 12
The impacts of active retail investor attention.

	Panel A — Retail Investor Attention			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
	Mean Daily Trading Volume in Millions			
ASVI	**0.2701*****	**0.2454*****	**0.2289*****	**0.2170*****
	8.52	*8.00*	*7.60*	*7.33*
Interaction	**−0.9039*****	**−0.8210*****	**−0.7238*****	**−0.7571*****
	−4.76	*−4.85*	*−4.72*	*−4.86*
Controls	YES	YES	YES	YES
F-Test	**68.95*****	**64.48*****	**58.29*****	**52.49*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.6333	0.6824	0.7104	0.7283
	Stock Returns Adjusted with Equal-Weighted Market Return			
ASVI	−0.0001	**0.0005*****	**0.0012*****	**0.0024*****
	−0.69	*3.98*	*7.31*	*11.95*
Interaction	**−0.0094*****	**−0.0052*****	−0.0028	**−0.0079*****
	−3.97	*−2.59*	*−0.80*	*−2.68*
Controls	YES	YES	YES	YES
F-Test	**15.36*****	**8.14*****	**36.82*****	**73.86*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.0131	0.0638	0.0682	0.0867
	Annualized Stock Return Volatilities			
ASVI	**0.4058*****	**0.3454*****	**0.2724*****	**0.2151*****
	16.88	*14.87*	*12.18*	*9.82*
Interaction	**−2.8190*****	**−2.1617*****	**−1.9058*****	**−2.2004*****
	−4.25	*−3.53*	*−2.84*	*−3.96*
Controls	YES	YES	YES	YES
F-Test	**143.20*****	**111.43*****	**75.08*****	**52.29*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.2176	0.2814	0.3145	0.3420
	Panel B — Change in Retail Investor Attention			
	1-Week Window	2-Week Window	3-Week Window	4-Week Window
	(1)	(2)	(2)	(4)
	Mean Daily Trading Volume in Millions			
Δ ASVI	**0.0268*****	**0.0205*****	**0.0168*****	**0.0144*****
	10.24	*10.96*	*11.09*	*10.75*
Interaction	**−0.2610*****	**−0.3328*****	**−0.3713*****	**−0.3547*****
	−2.37	*−3.55*	*−4.52*	*−4.43*
Controls	YES	YES	YES	YES
F-Test	**53.50*****	**60.49*****	**61.90*****	**58.43*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.6306	0.6800	0.7082	0.7263
	Stock Returns Adjusted with Equal-Weighted Market Return			
Δ ASVI	**−0.0003*****	**−0.0005*****	**−0.0009*****	**−0.0004*****
	−5.87	*−8.09*	*−12.35*	*−6.16*
Interaction	**0.0188*****	**0.0128*****	**0.0303*****	**0.0159*****
	4.62	*4.25*	*5.00*	*3.38*
Controls	YES	YES	YES	YES
F-Test	**18.28*****	**34.30*****	**83.51*****	**19.13*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.0135	0.0639	0.0685	0.0860
	Annualized Stock Return Volatilities			
Δ ASVI	**0.0831*****	**0.0745*****	**0.0628*****	**0.0611*****
	6.86	*9.74*	*9.93*	*11.44*
Interaction	**−1.6225*****	**−2.2278*****	**−2.4535*****	**−2.2241*****
	−2.56	*−4.45*	*−5.58*	*−5.48*
Controls	YES	YES	YES	YES
F-Test	**23.64*****	**48.79*****	**56.25*****	**70.92*****
N	277,650	277,650	277,650	277,650
Adj. R^2	0.2163	0.2805	0.3140	0.3416

This table reports regression results of OLS Models (1) and (2) where the dependent variables are mean daily trading volume, stock returns adjusted with equal-weighted market returns, and annualized stock return volatilities over the subsequent 1-week, 2-week, 3-week, and 4-week windows. Panel A presents results for OLS Model (1) and Panel B for OLS Model (2). Only results for variables of interest are reported. Stock and year fixed effects are controlled in the estimates. Standard errors are clustered by stock. F-test is against the null hypothesis that the coefficients on the key variable and the interaction term are jointly zero. T-statistics are reported in *Italics*. *, **, *** denote the significance level of 10%, 5%, and 1%, respectively.

Moreover, Panel B of Table 12 shows that the interaction term between the change in ASVI and 1-week stock return impacts mean daily trading volume, adjusted stock returns, and annualized stock return volatilities over the subsequent four windows, significant at the 1% level in all estimates. Again, this evidence suggests that recent stock returns influence the effect of retail

investor attention on stocks. In unreported tests, we find consistent results when active retail investor attention is measured with GSVI.

6. Conclusion

Utilizing weekly measures of retail investor base constructed with Robinhood investor data, and weekly measures of active retail investor attention constructed using the Google Search Volume Index for tickers, we find that active retail investor attention impacts retail investor base and therefore, stock demand and stock returns. These results provide support to the vast literature on active retail investor attention that assumes a positive correlation between such attention and retail investor base and a positive correlation between retail investor base and stock demand.

Our results show that (the change in) active retail investor attention is positively related to the size of (the change in) retail investor base in the following four weeks, indicating that attention increases retail investor base. These results are not driven by the endogeneity between active retail investor attention and retail investor base. Our results also show that retail investor base and the change in retail investor base are both positively related to mean daily trading volume and stock returns in the subsequent 1-week, 2-week, 3-week, and 4-week windows. This evidence supports the buying pressure hypothesis and the assumption that retail investor base positively affects stock demand and valuation. In contrast to the conjecture in previous studies, these effects are economically more significant for larger stocks than for smaller stocks.

We find the above relationships influenced by recent stock returns. With a higher level of active retail investor attention, winner stocks attract more retail investors than loser stocks, likely driven by winner chasing of retail investors. In contrast, with a temporary increase in active retail investor attention, loser stocks attract more retail investors in the subsequent weeks, likely driven by retail investors "buying the dip". These results suggest that active retail investor attention play a significant role in the trading behavior of retail investors. In addition, our results show that retail investor base and the increase in retail investor base contribute to higher annualized stock return volatilities in the subsequent four weeks, especially for winner stocks.

Our results still leave unanswered questions for future research. Robinhood investors might be different from other retail investors and their trading behavior might be more heavily impacted by attention. Due to data availability, it is unclear whether retail investors' trading behavior is different before the zero-commission era that starts in year 2015 and after the COVID-19 pandemic that starts in year 2020. The role of stock size in the interaction between the disposition effect and the winner-chasing behavior of retail investors also calls for future research.

CRediT authorship contribution statement

Zhongdong Chen: Conceptualization, Methodology, Data curation, Formal analysis Writing – original draft. **Karen Ann Craig:** Conceptualization, Methodology, Data curation, Writing – review & editing.

Appendix

See Table A.1 given here.

Table A.1
Summary of variables.

Variables	Description
Investors	The size of retail investor base in week t, proxied by two measures. The first measure, maximum retail investor base, is the maximum number of Robinhood investors for a stock in week t scaled by the maximum number of retail investors over the entire sample period. The second measure, average retail investor base, is the average daily number of Robinhood investors for a stock in week t scaled by the maximum number of retail investors over the entire sample period.
ΔInvestors	The change in retail investor base in week t, relative to week $t-1$.
Attention	The level of retail investor attention in week $t-1$, measured with GSVI and ASVI. GSVI is the natural logarithm of the Google Search Volume Index for tickers. ASVI is the natural logarithm of the average Google Search Volume Index for three search terms including "ticker", "ticker+stock", and"ticker+price".
ΔAttention	The change in retail investor attention in week $t-1$, relative to week $t-2$.
Ret_1W	Stock return in week $t-1$, adjusted with equal-weighted market return.
Ret_1M	Stock return in the prior 1-month period ending in week $t-1$, adjusted with equal-weighted market return.
Ret_3M	Stock return in the prior 3-month period ending in week $t-1$, adjusted with equal-weighted market return.
Volatility	Annualized stock return volatility in the prior year, calculated with daily stock returns in the prior year.
Illiquidity	Stock illiquidity, measured with zero return day ratio in the prior year. This ratio is calculated as the number of zero return trading days in the prior year divided by the total number of trading days in the prior year.
Price	Price level, calculated as one divided by stock price in week $t-1$.
BTM	Book-to-market ratio, calculated as book value per share in the prior year divided by stock price in week $t-1$.
Size	The natural logarithm of market capitalization in week $t-1$.
Institutional	Institutional stock ownership in the prior quarter.
Analyst	The natural logarithm of the number of analysts following in the prior quarter.
Buy	The natural logarithm of the number of buy recommendations in the prior quarter.
Advertising	The natural logarithm of annual advertising expenses in the prior year.
Earnings	An indicator variable that is equal to 1 if there is an earnings announcement in week $t-1$, and 0 otherwise.
Jump	An indicator variable that is equal to 1 if there is a stock price jump in week $t-1$, and 0 otherwise. Price jumps are a measure for information shock.
Volume	The ratio between mean daily trading volume in week $t-1$ and mean daily trading volume in week $t-2$.

References

Albuquerque, R., Koskinen, Y., Yang, S., Zhang, C., 2020. Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. Rev. Corp. Finance Stud. 9, 593–621.

Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. Rev. Financ. Stud. 28, 33–72.

Aouadi, A., Arouri, M., Teulon, F., 2013. Investor attention and stock market activity: Evidence from France. Econ. Model. 35, 674–681.

Ashour, S., Hao, G.Q., Harper, A., 2023. Investor sentiment, style investing, and momentum. J. Financ. Mark. 62, 100755.

Bae, K., Wang, W., 2012. What's in a China name? A test of investor attention hypothesis. Financ. Manag. 41, 429–455.

Bailey, W., Kumar, A., Ng, D., 2011. Behavioral biases of mutual fund investors. J. Financ. Econ. 102, 1–27.

Bank, M., Larch, M., Peter, G., 2011. Google search volume and its influence on liquidity and returns of German stocks. Financial Mark. Portfolio Manag. 25, 239–264.

Barber, B.M., Huang, X., Odean, T., Schwarz, C., 2022. Attention induced trading and returns: Evidence from Robinhood users. J. Finance 77, 3141–3190.

Barber, B.M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21, 785–818.

Ben-Rephael, A., Da, Z., Israelsen, R.D., 2017. It depends on where you search: A comparison of institutional and retail attention. Rev. Financ. Stud. 30, 3009–3047.

Bijl, L., Kringhaug, G., Molnár, P., Sandvik, E., 2016. Google searches and stock returns. Int. Rev. Financ. Anal. 45, 150–156.

Birru, J., 2015. Confusion of confusions: A test of the disposition effect and momentum. Rev. Financ. Stud. 28, 1849–1873.

Chai, D., Dai, M., Gharghori, P., Hong, B., 2021. Internet search intensity and its relation with trading activity and stock returns. Int. Rev. Finance 21, 282–311.

Chemmanur, T.J., Yan, A., 2019. Advertising, attention, and stock returns. Q. J. Finance 9, 1950009.

Chen, T., 2017. Investor attention and global stock returns. J. Behav. Finance 18, 358–372.

Chen, Z., 2021. Investor attention and market correction. Rev. Behav. Finance 13, 386–409.

Chen, H., Noronha, G., Singhal, V., 2004. The price response to S & P 500 Index additions and deletions: Evidence of asymmetry and a new explanation. J. Finance 59, 1901–1929.

Chen, R., Qian, Q., Jin, C., Xu, M., Song, Q., 2020. Investor attention on internet financial markets. Finance Res. Lett. 36, 101421.

Chen, Z., Schmidt, A., Wang, J., 2021. Retail investor risk-seeking, attention, and the January effect. J. Behav. Exp. Finance 30, 100511.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. J. Finance 66, 1461–1499.

Da, Z., Engelberg, J., Gao, P., 2015. The sum of all FEARS investor sentiment and asset prices. Rev. Financ. Stud. 28, 1–32.

Dang, V.A., Michayluk, D., Pham, T.P., 2018. The curious case of changes in trading dynamics: When firms switch from NYSE to NASDAQ. J. Financial Mark. 41, 17–35.

Desagre, C., D'Hondt, C., 2021. Googlization and retail trading activity. J. Behav. Exp. Finance 29, 100453.

Dimpfl, T., Jank, S., 2016. Can internet search queries help to predict stock market volatility? Eur. Financial Manag. 22, 171–192.

Ding, R., Hou, W., 2015. Retail investor attention and stock liquidity. J. Int. Financ. Mark. Inst. Money 37, 12–26.

Ding, R., Zhou, H., Li, Y., 2020. Social media, financial reporting opacity, and return comovement: Evidence from seeking alpha. J. Financ. Mark. 50, 100511.

Drake, M.S., Jennings, J., Roulstone, D.T., Thornock, J.R., 2017. The comovement of investor attention. Manage. Sci. 63, 2773–3145.

Dreman, D., Johnson, S., MacGregor, D., Slovic, P., 2001. A report on the 2001 investor sentiment survey. J. Psychol. Financ. Mark. 2, 126–134.

Eaton, G.W., Green, C.T., Roseman, B.S., Wu, Y., 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. J. Financ. Econ. 146, 502–528.

Fama, E.F., French, K.R., 2018. Choosing factors. J. Financ. Econ. 128 (2), 234–252.

Gervais, S., Kaniel, R., Mingelgrin, D.H., 2001. The high-volume return premium. J. Finance 56, 877–919.

Grullon, G., Kanatas, G., Weston, J.P., 2004. Advertising, breadth of ownership, and liquidity. Rev. Financ. Stud. 17, 439–461.

Haruvy, E., Lahav, Y., Noussair, C., 2007. Traders' expectations in asset markets: Experimental evidence. Amer. Econ. Rev. 97, 1901–1920.

Hervé, F., Zouaoui, M., Belvaux, B., 2019. Noise traders and smart money: Evidence from online searches. Econ. Model. 83, 141–149.

Herwartz, H., Xu, F., 2022. Structural transmissions among investor attention, stock market volatility and trading volumes. Eur. Financ. Manag. 28, 260–279.

Heyman, D., Lescrauwaet, M., Stieperaere, H., 2019. Investor attention and short-term reversals. Finance Res. Lett. 29, 1–6.

Hou, K., 2007. Industry information diffusion and the lead–lag effect in stock returns. Rev. Financ. Stud. 20, 1113–1138.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Rev. Financ. Stud. 28 (3), 650–705.

Israeli, D., Kasznik, R., Sridharan, S.A., 2022. Unexpected distractions and investor attention to corporate announcements. Rev. Account. Stud. 27, 477–518.

Jankensgård, H., Vilhelmsson, A., 2018. The shareholder base hypothesis of stock return volatility: Empirical evidence. Financ. Manag. 47, 55–79.

Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise—A swap variance approach. J. Econometrics 144, 352–370.

Jiang, G.J., Zhu, K.X., 2017. Information shocks and short-term market underreaction. J. Financ. Econ. 124, 43–64.

Joseph, K., Wintoki, B.M., Zhang, Z., 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. Int. J. Forecast. 27, 1116–1127.

Kaniel, R., Ozoguz, A., Starks, L., 2012. The high volume return premium: Cross-country evidence. J. Financ. Econ. 103, 255–279.

Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. J. Finance 63, 273–310.

Kao, L., Chen, A., Lu, C., 2022. Retail investor attention and IPO prices with a pre-IPO market. Int. Rev. Econ. Finance 82, 416–432.

Kim, N., Lučivjanská, K., Molnár, P., Villa, R., 2019. Google searches and stock market activity: Evidence from Norway. Finance Res. Lett. 28, 208–220.

Kostopouls, D., Meyer, S., Uhr, C., 2020. Google search volume and individual investor trading. J. Financial Mark. 49, 100544.

Kumar, A., Lee, C.M.C., 2006. Retail investor sentiment and return comovements. J. Finance 61, 2451–2486.

Kupfer, A., Schmidt, M.G., 2022. In search of retail investors: The effect of retail investor attention on odd lot trades. J. Empir. Financ. 62, 315–326.

Lee, C.M.C., So, E.C., 2017. Uncovering expected returns: Information in analyst coverage proxies. J. Financ. Econ. 124, 331–348.

Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction cost. Rev. Financ. Stud. 12, 1113–1141.

Liaukonyte, J., Zaldokas, A., 2022. Background noise? TV advertising affects real-time investor behavior. Manage. Sci. 68, 2377–3174.

Lo, A.W., Wang, J., 2006. Trading volume: Implications of an intertemporal capital asset pricing model. J. Finance 61, 2805–2840.

Lou, D., 2014. Attracting investor attention through advertising. Rev. Financ. Stud. 27, 1797–1829.

Madsen, J., Niessner, M., 2019. Is investor attention for sale? The role of advertising in financial markets. J. Account. Res. 57, 763–795.

Massa, M., Žaldokas, A., 2014. Investor base and corporate borrowing: Evidence from international bonds. J. Int. Econ. 92, 95–110.

Mayer, E.J., 2021. Advertising, investor attention, and stock prices: Evidence from a natural experiment. Financ. Manag. 50, 281–314.

McCrank, J., 2021. Factbox: The U.S. retail trading frenzy in numbers. Reuters. https://www.reuters.com/article/us-retail-trading-numbers/factbox-the-u-s-retail-trading-frenzy-in-numbers-idUSKBN29Y2PW.

Meng, J., Weng, X., 2018. Can prospect theory explain the disposition effect? A new perspective on reference points. Manage. Sci. 64, 2973–3468.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. J. Finance 42, 483–510.

Meshcheryakov, A., Winters, D.B., 2022. Retail investor attention and the limit order book: Intraday analysis of attention-based trading. Int. Rev. Financ. Anal. 81, 101627.

Moore, J., 2020. Glamour among value: P/E ratios and value investor attention. Financ. Manag. 49, 673–706.

Moss, A., Naughton, J.P., Wang, C., 2020. The Irrelevance of ESG Disclosure to Retail Investors: Evidence from Robinhood. Working Paper, University of Iowa.

Odean, T., 1998. Are investors reluctant to realize their losses? J. Finance 53, 1775–1798.

Pagano, M.S., Sedunov, J., Velthuis, R., 2021. How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality. Finance Res. Lett. 43, 101946.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. J. Financ. Econ. 80, 563–602.

Poterba, J.M., Weisbenner, S.J., 2001. Capital gains tax rules, tax loss trading, and turn-of-the-year returns. J. Finance 56, 353–368.

Rakowski, D., Shirley, S.E., Stark, J.R., 2021. Twitter activity, investor attention, and the diffusion of information. Financ. Manag. 50, 3–46.

Rau, H.A., 2015. The disposition effect in team investment decisions: Experimental evidence. J. Bank. Financ. 61, 272–282.

Robinhood, https://investors.robinhood.com/overview/default.aspx, (last viewed September 20, 2022).

Seasholes, M.S., Wu, G., 2007. Predictable behavior, profits, and attention. J. Empir. Finance 14, 590–610.

Shohfi, T., Simaan, M., 2021. Buy the Dip. Working paper, Rensselaer Polytechnic Institute.

Solomon, D.H., Soltes, E., Sosyura, D., 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. J. Financ. Econ. 113, 53–72.
Swamy, V., Dharani, M., 2019. Investor attention using the google search volume index – Impact on stock returns. Rev. Behav. Finance 11, 56–70.
Takeda, F., Wakao, T., 2014. Google search intensity and its relationship with returns and trading volume of Japanese stocks. Pac.-Basin Finance J. 27, 1–18.
Tang, D.Y., Zhang, Y., 2020. Do shareholders benefit from green bonds? J. Corporate Finance 61, 101427.
Vlastakis, N., Markellos, R.N., 2012. Information demand and stock market volatility. J. Bank. Financ. 36, 1808–1821.
Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. J. Bank. Finance 41, 17–35.
Welch, I., 2022. The wisdom of the Robinhood crowd. J. Finance 77, 1489–1527.
Yang, D., Ma, T., Wang, Y., Wang, G., 2021. Does investor attention affect stock trading and returns? Evidence from publicly listed firms in China. J. Behav. Finance 22, 368–381.
Yezegel, A., 2015. Why do analysts revise their stock recommendations after earnings announcements? J. Account. Econ. 59, 163–181.
Ying, Q., Kong, D., Luo, D., 2015. Investor attention, institutional ownership, and stock returns: Empirical evidence from China. Emerg. Mark. Finance Trade 51, 672–685.
Zhang, B., Wang, Y., 2015. Limited attention of individual investors and stock performance: Evidence from the ChiNext market. Econ. Model. 50, 94–104.