Which Attention Matters?

Chuyi Yang[1]
Xiamen University

Using novel user-activity data from a leading Chinese brokerage, I decompose investor attention into four distinct channels: community forums (unsophisticated retail attention), research reports (sophisticated retail and institutional attention), news, and announcements. Community attention is positively correlated with contemporaneous returns but negatively predicts future returns, consistent with noise trader overreaction. In sharp contrast, research report attention positively predicts returns without reversal, reflecting fundamental information incorporation. News attention exhibits positive predictability followed by partial reversal, while announcement attention shows delayed positive effects. All attention effects are conditional on market sentiment. These findings highlight that aggregating attention sources masks opposing investor behaviors.

Keywords: Investor Attention; Retail Attention; Institutional Attention; Sentiment
JEL Codes: G11, G12, G14, G41

[1] Chuyi Yang: Department of Finance, School of Economics and The Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, 422 South Siming Road, Xiamen 361005, Fujian, China, email: chuyiyang@xmu.edu.cn. Please send all correspondence to Chuyi Yang (corresponding author). Declarations of interest: none. Funding: This work was supported by National Natural Science Foundation of China [grant number 72202191]; the Fundamental Research Funds for the Central Universities [grant number 2072021146].

I. Introduction

Retail investors[2] are at center stage in China's stock market, constituting the largest population of individual investors globally. According to the China Securities Depository and Clearing Corporation (CSDC) Statistical Yearbook, the number of retail investor accounts in the A-share market reached approximately 0.2 billion by the end of 2023. The Chinese stock market also exhibits a clear hierarchy of investor sophistication, which is strongly correlated with wealth and investor type. Small retail investors are momentum traders who predict returns incorrectly, while institutional investors and large retail investors are contrarian investors, who predict returns correctly (Tan, Zhang and Zhang, 2024). This dichotomy is rooted in information processing: smaller retail investors often fail to incorporate public information effectively, whereas large retail investors and institutions are likely to process firm-level fundamental information (Jones et al., 2025).

Investor attention is a fundamental driver of retail trading behavior (Jiang et al., 2022; Tan, Zhang, and Zhang, 2024) and plays a crucial role in asset pricing. However, accurately measuring investor attention remains a central challenge in the literature. Existing studies largely rely on indirect proxies[3], primarily internet search volume from engines like Google or Baidu. While these

[2] Retail investors or individual investors register their accounts with personal identification cards. Institutional investors register their accounts with government-issued permits and licenses.

[3] Alternative proxies for investor attention include news, unusual trading volume, and extreme returns (Barber and Odean, 2008), price-limit events (Seasholes and Wu, 2007), and daily investor online account logins (Sicherman et al., 2016). Investor attention is affected by firm advertising (Grullon et al., 2004; Chemmanur and Yan, 2009; Lou, 2014; Focke, Ruenzi, and Ungeheuer, 2020), high trading volume (Gervais, Kaniel and Mingelgrin, 2001) and market state (Hou, Xiong and Peng, 2009).

search-based measures have yielded fruitful insights, they are indirect proxies of investor attention or cannot fully capture attention towards a stock. They often fail to distinguish genuine financial interest from general information-seeking. For instance, a search for "Apple" may not pertain to the stock but to the consumer product, introducing measurement error. Some studies have progressed by using more direct proxies. Using number of posts on EastMoney investor forum as proxy for investor attention, Jiang et al. (2022) document that high attention leads to overreaction and higher anomaly returns.

This paper contributes to the investor attention literature by leveraging user-behavior data from EastMoney, a leading Chinese brokerage with over 17 million monthly active users[4]. First, current literature on attention largely focuses on one or two dimensions of attention from different data sources, making comparisons across platforms less direct. In this paper, the measures are derived from direct user interactions within a financial platform, mitigating the ambiguity inherent in search-based proxies and providing a more precise gauge of investor-specific attention. Specifically, direct measures of attention are based on investors' reading and commenting behavior from EastMoney. Second, this paper decomposes investor attention into distinct, behaviorally meaningful types: attention to community forums, news, corporate announcements, and research reports. This allows us to differentiate between attention influenced by investor behavior on investor forum (community) and attention driven by formal information arrival (news, announcements, reports). I use community attention as a proxy for "unsophisticated attention",

[4] The number of monthly active users on the EastMoney app as of March 2025 is 17.2186 million.

dominated by unsophisticated retail investors; whereas research reports attention as a proxy for "sophisticated attention", dominated by institutional and sophisticated retail investors.

Building on this novel dataset, I construct four distinct attention indexes to capture different facets of investor engagement: *Community Attention Index* is measured by number of posts and replies in a firm-specific stock message board on Guba Eastmoney[5]. This index primarily reflects the attention of noise traders, dominated by unsophisticated retail investors. *Research Report Attention Index* is derived from number of reads and comments on professional research reports, reflecting attention of sophisticated traders dominated by sophisticated retail investors and institutional investors. *News Attention Index* is based on the number of news article reads and associated comments for a given stock. *Announcement Attention Index* is calculated from the number of reads and comments on corporate announcements. The separation of attention sources reveals a fundamental divergence in how these attention types relate to returns. Contemporaneously, attentions towards community, research reports, news and announcements are also positively correlated with returns. The predictive power of these measures diverges. *Community Attention Index* negatively predicts next-day returns, consistent with retail-driven overreaction that subsequently reverses. Conversely, attention to News, Announcements, and Research Reports positively predicts future returns. This positive predictability for *Research Reports Attention Index* persists for 5 trading days without reversal, suggesting it captures

[5] Guba EastMoney, established in 2006, is the largest and most active online stock forum in China. For each stock, there is a firm-specific sub-forum in Guba EastMoney, where investors can post their views and reply other's views on a particular stock.

fundamental, value-relevant information. Specifically, a one percentage increase in *Community Attention Index* leads to 4.6 bps decrease of stock return for the next trading days, and the cumulative negative effect of 26.9 bps over the next 10 trading days. In sharp contrast, a one percentage increase in *Research Report Index* predicts 21.3 bps increase of stock return next trading day, and cumulative increase of 28.1 bps in return for the next 5 trading days. A similar sustained effect is observed for *Announcement Attention Index* with delayed response, while the predictive power of *News Attention Index* exhibits reversal, indicating correction of initial overreaction, both of which are more transient informational component than research reports.

While the literature on investor attention is well-established, the research largely focuses on retail[6] and institutional attention. This paper instead proposes the perspective from sophisticated attention (from institutional investors and sophisticated retail investors) and unsophisticated attention (from unsophisticated retail investors). Since the U.S. and Europe markets are dominated by institutional investors, this study contributes to the literature by examining attention in the Chinese stock market, the second-largest stock market in the world where tens of millions of retail investors play a dominant role. Retail investors contribute to about 85%[7] of the daily trading volume on the exchange based on the annual report of the Shanghai Stock Exchange in 2016 (Jones

[6] Studies on retail investors include Barber and Odean (2001, 2002, 2008), Barber, Odean, and Zhu (2009), Kaniel, Saar, and Titman (2008), Kelley and Tetlock (2013), Boehmer, Jones, Zhang, and Zhang (2021), Barber, Huang, Odean, and Schwarz (2022), Eaton, Green, Roseman, and Wu (2022), and Welch (2022).

[7] While this share has declined in recent years due to growing institutional participation, retail investors still represented a substantial majority, contributing roughly 60% of trading volume as of the end of 2022. This persistent and significant market share underscores the critical importance of understanding retail investor behavior for analyzing stock return in the Chinese market.

et al., 2025), but they are generally characterized by low level of financial literacy (Song, 2020; Titman, Wei and Zhang, 2022; Tan, Zhang and Zhang, 2024). The findings in the paper provide direct, behavioral evidence for the profound heterogeneity within the Chinese investor base, as suggested by Ng and Wu (2007) and Tan, Zhang, and Zhang (2024). I demonstrate that the attention of unsophisticated retail investors (proxied by community attention) with low level financial literacy and behavioral biases, negatively predicts future returns, consistent with attention-driven behavioral biases. In contrast, the attention of sophisticated retail and institutional investors (proxied by research report attention) positively predicts returns, underscoring their ability to process fundamental information. Thus, this paper not only confirms the existence of a sophistication hierarchy but also identifies differential attention allocation as a key mechanism driving it.

Trading behavior of retail investor is consistent with theories of noise trader in De Long et al. (1990a). According to the limits-to-arbitrage literature (Shleifer and Summers, 1990; Shleifer and Vishny, 1997; De Long et al., 1990a, 1990b), liquidity is provided by noise traders who often possess wrong beliefs or are overconfident about their information. It is found that *Community Attention Index* predicts small and middle trade order imbalance positively while predicts large and extra-large trade order imbalance negatively. In contrast, *Research Report Attention Index* predicts small and middle trade order imbalance negatively, but predicts large and extra-large trade order imbalance positively. Therefore, sophisticated trader attention captured by *Research Report Attention Index* positively predicts permanent price changes, consistent with deliberate attention allocation and information acquisition. Rational inattention models suggest that investors'

attention allocation decisions are outcomes of investor optimization, depending on their attention capacity and the nature of news (Sims, 2003, 2006; Peng, 2005; Peng and Xiong, 2006; Mondria, 2010; Van Nieuwerburgh and Veldkamp, 2010; Gondhi, 2021). In sharp contrast, unsophisticated attention from retail investors proxied by *Community Attention Index* generates negative pressure associated with transient price effects, consistent with salience-based attention. Retail investors are naturally drawn to salient stimuli and can be automatic and involuntary, which distract retail investors from deliberate attention and result in suboptimal decisions (Bordalo, Gennaioli, and Shleifer, 2012, 2013a, b, 2020). I find that sophisticated traders tend to buy from and sell to noise traders, systematically profiting from their behavioral biases.

If uninformed noise traders base their trading decisions on sentiment and risk-averse arbitrageurs encounter limits to arbitrage, sentiment changes will lead to more noise trading and greater mispricing (De Long et al., 1990a). To test the interaction between attention and sentiment, I construct a daily sentiment measure based on the frequency of positive and negative words in Guba EastMoney posts. During high-sentiment periods, elevated *Community Attention Index* predicts return positively that subsequently reverse (T: +33.8 bps; T+1: -3.1 bps)—a pattern consistent with sentiment-driven, temporary buying pressure. During low sentiment period, *Community Attention Index* is negatively correlated with stock return without subsequent reversal (T: -37.5 bps; T+1: -10.4 bps). Specifically, a one percentage increase in *Community Attention Index* leads to -3.1 bps decrease of stock return on the next trading days during high-sentiment period, and 10.4 bps decrease of stock return for the next trading days during low-sentiment days. In sharp contrast, *Research Report Attention Index*, proxy for sophisticated retail and institutional

attention, predicts return positively on both high and low sentiment days. A one percentage increase in *Research Report Attention Index* leads to 3.6 bps increase of stock return for the next trading days during low-sentiment periods and 5.7 bps during high-sentiment periods.

This paper provides further evidence on the attention literature that retail and institutional investor pay attention differently. Uninformed investors are typical users of the Baidu search engine (Guo, Jia and Sun, 2023). Existing literature[8] measures institutional investor attention based on news searching and news reading activity on Bloomberg terminals, while retail attention is measured by Google search activity (Da, Engelberg, and Gao, 2011; Ben-Rephael, Da, and Israelsen, 2017). It is found that institutional attention, leading retail attention, responds more quickly to major news events and facilitates permanent price adjustment, with firm-specific news as the most important driver of abnormal institutional attention in the U.S. stock market (Ben-Rephael, Da, and Israelsen, 2017). In this paper, the dichotomy between institutional and retail attention, documented in U.S. markets, is powerfully examined in the Chinese stock market where research reports are the most important driver of sophisticated attention.

The central thesis in the limited attention literature[9] is that attention is scarce and requires effort (Kahneman, 1973).[10] This paper contributes to the literature on investor attention and stock

[8] A significant strand of this research exploits quasi-natural experiments, such as Google's withdrawal from China, to identify causal effects (Xu, Xuan, and Zheng, 2021; Wang, Yu, and Zhang, 2023).

[9] Existing evidence includes the effect of investor attention on return co-movement (Huang et al., 2019), trading activity (Dorn, Dorn, and Sengmueller, 2015; Gao and Lin, 2015; Peress and Schmidt, 2020), liquidity, volatility, and price reversal (Peress and Schmidt, 2020), investment performance (Bui, Huang, Lin, and Lin, 2021), and anomaly return (Jiang, Liu, Peng, and Wang, 2022).

[10] Limited attention would also influence investor perceptions, resulting in neglecting information and underreacting to news (Hirshleifer and Teoh, 2003; Peng, 2005; Peng and Xiong, 2006; Hirshleifer, Lim, and Teoh, 2009).

return predictability by moving beyond aggregate or sentiment-based measures to dissect the distinct informational roles of different attention channels. Numerous studies have shown that limited attention affects market returns (Li and Yu, 2012; Yuan, 2015; Chen et al., 2021; Da et al., 2025), including through attention spillover effects (Chen et al., 2023).

However, existing literature largely treats social media or message boards[11] as a monolithic source of "uninformed" or "retail" sentiment. It sometimes conflates the attention-grabbing nature of the platform with the intrinsic value of the information type being discussed. This is where this paper makes its central contribution by arguing that attention itself is multifaceted. By decomposing trading platform activities into discrete channels—community forums, news, announcements, and research reports—I can isolate attention driven by peer sentiment from attention driven by the arrival of fundamental information. This framework reconciles seemingly contradictory findings in the prior literature by showing that the source of attention determines its predictive power and persistence, not merely the platform on which it occurs.

Section II discusses data and measurement. Section III presents baseline results. Section IV provides evidence on attention and sentiment. Section V concludes.

II. Data and Measurement

The sample includes all A-share stocks (excluding ST stocks) listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange from January 1st, 2017 to March 31st, 2024. Stock

[11] Researchers have devoted considerable effort in examining the potential effect of stock message boards on financial market, including Antweiler and Frank (2004), Das and Chen (2007), Sabherwal et al. (2011), Chen et al. (2014), Bartov, Faurel, and Mohanram (2018), Gu and Kurov (2020), Booker, Curtis, and Richardson (2023).

returns and accounting information are from the Chinese Stock Market and Accounting Research (CSMAR) database.

i. Attention Measures

Attention measures are constructed based on investors' posts and replies in individual stock forums (*Community Attention Index*), investors' reads and comments of news (*News Attention Index*), announcements (*Announcement Attention Index*) and research reports (*Research Report Attention Index*) on EastMoney app. Application of message boards in Chinese stock markets are explored in Hong et al. (2014), Jiang, Liu and Yang (2018), Chang et al. (2022), Huang, Qiu, and Wu (2016) and Li, Shen and Wang (2018), who use China's Internet stock message board Guba EastMoney. Jiang et al. (2022) constructs the stock-level attention measure based on number of posts at Guba EastMoney. Because the EastMoney app integrates these information channels with trading functionality, the measures capture attention from investors who are poised to act, providing a more precise and behaviorally-grounded test of how different information types influence trading and returns. While these studies establish the significant influence of this platform, they predominantly rely on a single, aggregate dimension of attention—typically post volume. Different from prior literature that largely focuses on one dimension of attention, this paper further separates attention by its sources.

I construct the abnormal investor attention measure following Da et al. (2011). The adjusted measures of attention are as follows:

$$\log(\text{Abnormal Attention}_t)$$
$$= \log(\text{Attention}_t) - \log([\text{Med}(\text{Attention}_{t-1}, \text{Attention}_{t-2}, \dots \text{Attention}_{t-30})])$$

where $\log(\text{Attention}_t)$ is the logarithm of attention during day t, and [$\text{Med}(\text{Attention}_{t-1}$, Attention_{t-2},…Attention_{t-30})] is the logarithm of the median value of SVI during the prior 30 trading days. The median over a longer time window captures the "normal" level of attention in a way that is robust to recent jumps. $\log(\text{Abnormal Attention}_t)$ removes time trends and other low-frequency seasonality. A large positive $\log(\text{Abnormal Attention}_t)$ clearly represents a surge in investor attention and can be compared across stocks in the cross-section. Log(Abnormal attention) measures include *Community Attention Index*, *News Attention Index*, *Announcement Attention Index* and *Research Report Attention Index*.

ii. Buy-Sell Imbalance

We follow Kumar and Lee (2006) to construct Buy-Sell Imbalance (BSI) of stocks as proxy for retail sentiment. A given stock-level BSI indicates whether, at an aggregate level, investors are net buyers (stock BSI > 0, that is, a positive change in their aggregate stock sentiment) or net sellers (stock BSI < 0, that is, a negative change in their aggregate stock sentiment) of a given stock on a given day. In other words, the stock-level BSI measure is a directional indicator of net retail/institutional demand for that stock in a given day. The trading activities information is obtained from EastMoney moneyflow. Specifically, EastMoney applies the following classifications: small trade includes transaction orders with fewer than 20,000 shares and a total value under CNY 40,000; middle trade includes transaction orders with larger than or equal to 20,000 shares and a total value of CNY 40,000, as well as smaller than 100,000 shares and a total value under CNY 200,000; large trade includes transaction orders with larger than or equal to 100,000 shares and a total value of CNY 200,000, as well as smaller than 500,000 shares and a

total value under CNY 1000,000; extra-large trade includes transaction orders with larger than or equal to 500,000 shares and a total value of CNY 1000,000. I use transaction orders with fewer than 100,000 shares and a total value under CNY 200,000 as proxy for unsophisticated retail investors' trading activities. I use transaction orders with larger than or equal to 100,000 shares and a total value of CNY 200,000 as proxy for sophisticated retail and institutional investors' trading activities. This is consistent with the classification in EastMoney that small trade and middle trade are proxy for unsophisticated retail trading, while large trade and extra-large trade are proxy for sophisticated retail and institutional trading. Two measures of BSI based on amount of transaction and volume of transaction are constructed as follows:

$$BSI_Amount_{it} = \frac{Buy_Amount_{it} - Sell_Amount_{it}}{Buy_Amount_{it} + Sell_Amount_{it}}$$

iii. Sentiment

Unlike formal media such as newspapers or large news websites, stock forum comments belong to informal media. The content of these posts primarily reflects the subjective opinions of investors, characterized by a high degree of randomness and irregularity. For example, stock forum comments may consist of just a few simple words, irrelevant emojis, or meaningless text expressions—all of which introduce noise that can affect the accuracy of sentiment analysis on these posts. To improve the accuracy of sentiment judgment, a supervised learning model is employed to classify the sentiment of stock forum posts. The application of supervised learning in this context involves the following steps: first, three sentiment categories are established—*positive*, *negative*, and *neutral*. Posts are manually labeled accordingly (e.g., *1* for positive, *-1* for negative,

and *0* for neutral). Second, a labeled dataset is compiled, with a portion automatically extracted to serve as training data. Third, a Support Vector Machine (SVM) algorithm is applied to the training dataset to learn and generate a classification model. Fourth, the trained model is then used to automatically predict the sentiment categories of new, unlabeled posts in the test dataset. This structured approach enhances the reliability of sentiment analysis in the noisy, unstructured environment of stock forums.

Based on the above-mentioned method, each stock forum post's sentiment is classified into positive, negative, or neutral, which is obtained from Chinese Research Data Services (CNRDS). The daily totals for each category within a listed company's forum are aggregated. Therefore, the following ratio describes the sentiment of each stock.

$$Pos_Neg_{it} = \frac{Pos_{it} - Neg_{it}}{Pos_{it} + Neg_{it}}$$

Pos_{it} is the number of positive comments for stock i on day t.

Neg_{it} is the number of negative comments for stock i on day t.

I include logarithm of market size (Size), earnings to price ratio and turnover as control variables to control for effect of size, value, and liquidity on stock return, consistent with Jones et al. (2025).

III. Results

Table I Panel A presents summary statistics and correlation of main variables. The *Community* and *Research Report Index* have slightly negative average values of -0.058 and -0.161, while the *News* and *Announcement Index* have slightly positive average values of 0.249 and 0.253. Stock Returns exhibit a mean of 0.032%, with extreme values ranging from -82.581% to 20.833%. The Buy-Sell

Imbalance (BSI) measures for Unsophisticated and Sophisticated investors show opposing average tendencies (0.018 and -0.067, respectively). Other firm-level characteristics include an average EP Ratio of 0.087 with high variability, a mean Turnover of 0.018, and the natural log of Market Value averaging 15.917. Panel B presents correlation of main variables. The pairwise correlation between all attention measures are positive. This is consistent with findings in Cookson et al. (2024) that attention is highly correlated across three platforms Twitter, StockTwits, and Seeking Alpha, but sentiment is not, after controlling for firm disclosures and news. Correlation between unsophisticated attention (proxied by Community Attention Index) and sophisticated attention (proxied by Research Report Attention Index) is 0.2812. The largest correlation 0.4167 is between news and announcements.

[Insert Table I Here]

i. Baseline Results

To investigate the effect of attention on stock returns, I run daily Fama and MacBeth (1973) regression. The t-statistics are robust to heteroskedasticity and autocorrelation. Tetlock (2011) uses similar method to study the effect of firm-specific news on stock returns. Gu and Kurov (2020) adopt the method to study the effect of Twitter sentiment on stock returns. Return momentum of 5 lags of return is controlled for, since return autocorrelation in conjunction with contemporaneous correlation of returns and attention can generate spurious evidence of lead-lag relation (Chordia and Swaminathan, 2000; Rapach et al., 2013). Studies on retail investors also adopt similar

approaches (Kelly & Tetlock, 2013, Jones et al., 2025). The dependent variable is the return on day *t* for stock *i*, where k = 0, 1, 2…11.

$$Ret_{i,t+k} = a0 \quad + \beta 1_{i,t} Attention_{i,t} + \beta 2_{i,t} \log_Size_{i,t} + \beta 3_{i,t} EP_ratio_{i,t} + \beta 4_{i,t} Turnover_{i,t}$$
$$+ \beta 5_{i,t} Ret_lag1_{i,t-1} + \beta 6_{i,t} Ret_lag2_{i,t-2} + \beta 7_{i,t} Ret_lag3_{i,t-3}$$
$$+ \beta 8_{i,t} Ret_lag4_{i,t-4} + \beta 9_{i,t} Ret_lag5_{i,t-5} + \varepsilon_{i,t+k}$$

Table II studies the effect of community and research attention indexes on stock return controlling for news and announcement attention indexes. *Community Attention Index* more accurately reflects the preferences of unsophisticated retail investors, while *Research Report Attention Index* more accurately reflects sophisticated retail and institutional investors (e.g. the index increases significantly during mid-year and quarterly reports). All four attention measures are positively correlated with stock return contemporaneously in Column (1). Specifically, one percentage increase in *Community Attention Index*, *Research Report Attention Index*, *News Attention Index*, *Announcement Attention Index*, and increases contemporaneous return by 3 bps, 13.7 bps, 14 bps, and 1.2 bps respectively.

In terms of return predictability on the next trading day, *Community Attention Index* negatively predicts stock return, while *News Attention Index*, *Announcement Attention Index*, and *Research Report Attention Index* positively predict stock return in Column (2). The different return predictability patterns of *Community Attention Index* and *Research Report Attention Index* suggest that unsophisticated traders and sophisticated traders pay attention differently. Unsophisticated attention, proxied by *Community Attention Index*, predicts return negatively up to 10 trading days, while sophisticated attention, proxied by *Research Report Attention Index*, predicts return

positively up to 5 trading days continuously without return reversal. Specifically, a one percentage increase in *Community Attention Index* leads to 4.6 bps decrease of stock return the next trading day, and the cumulative negative effect of 26.9 bps over the next 10 trading days (T+1 to T+10). The complete reversal of return effect by *Community Attention Index* is consistent with temporary upward price pressure caused by high investor attention of noise traders (De Long et al., 1990a), which leads to overreaction in the short run that subsequently reverses. This is also consistent with Barber and Odean (2002) that trading is hazardous to individual investor's wealth. Given the dominant trading behavior by retail investors[12] in the Chinese stock market, the influence of wealth could be substantial. In sharp contrast, a one percentage increase in *Research Report Attention Index* predicts 21.3 bps increase of stock return the next trading day, and cumulative increase of 28.1 bps in return for the next 5 trading days (T+1 to T+5). This suggests that attention towards *Research Report* contains information that is not incorporated into the stock price, as there is no return reversal. The results are consistent with existing evidence on retail and institutional trading in the China market. Tan, Zhang and Zhang (2024) find that small retail investors have low financial literacy, exhibit behavioral biases, and therefore negatively predict future returns, whereas large retail investors and institutions are capable of processing information and positively predict future returns.

[12] Based on proprietary data from one major stock exchange in China from January 2016 to June 2019, Jones et al. (2025) find that retail investors hold only approximately 20% of all shares while contribute to approximately 80% of total volumes, while institutions account for less than 20% of all shares and less than 20% of total volumes, block holders hold approximately 60% but barely trade.

Existing research consistently demonstrates that divergence in attention between retail and institutional investors translates into opposing return predictability. Da et al. (2025) use Google's daily Search Volume Index (SVI) as proxy for retail attention and Bloomberg's daily maximum readership for a given stock as proxy for institutional attention. They find that daily aggregate retail attention negatively predicts one-week-ahead market returns; while aggregate institutional attention, when measured before major news announcements, positively predicts future market-wide returns. Da et al. (2025) argue that institutional investors pay attention rationally while retail investors pay attention to stimuli using U.S. sample. This underscores that the relationship between attention and returns is fundamentally shaped by heterogeneous investor clienteles (Griffin, Harris and Topaloglu, 2003; Barber and Odean, 2008; Da, Engelberg, and Gao, 2011; Barber et al., 2022).

A one percentage increase in *News Attention Index* predicts 18.4 bps increase of stock return the next trading day, which subsequently partially reverses. The cumulative positive effect for the 2 trading days (T+1 to T+2) is 20.5 bps. The reversal effect from the next 3 to 7 trading days (T+3 to T+7) is -3.1 bps, suggesting that potential overreaction towards news reading in the short run that reverses in the longer term. Unlike *Community Attention Index*, there is no complete return reversal for *News Attention Index,* suggesting attention towards news contains relevant information about fundamentals.

The competition for attention is also illustrated during earnings announcement periods. Liu, Peng and Tang (2023) find that macroeconomic news crowds out retail investor attention to firm-specific earnings news by 49% while increases institutional investor attention to scheduled earnings announcements. Hirshleifer and Sheng (2022) also examine institutional and retail

attention when there is coincidence of macroeconomic news releases and individual firm earnings announcements. A one percentage increase in *Announcement Attention Index* is associated with 0.6 bps increase of stock return the next trading day. Delayed response of 0.4 bps and 0.3 bps happen at T+4 and T+5 trading days.

Cookson et al. (2024) argue that social media sentiment predicts positive next-day returns, but social media attention predicts negative next-day returns from three major platforms: Twitter, StockTwits, and Seeking Alpha. If attention contains no information but temporarily shifts the demand of stock, then return would fluctuate in the short run followed by return reversals. If attention conveys valuable information that is not already priced in the stock, then attention would affect stock return without return reversal. These suggests that *Community Attention Index* is more aligned with social media attention, which predicts negative next-day return. Higher attention temporarily shifts the demand of stock in the market dominated by retail investors. This is consistent with models of attention-induced trading and returns, which predicts that periods of intense buying will be followed by negative abnormal returns (Barber and Odean, 2008). Consistent with predictions of attention-based models of trading, Barber et al. (2022) use Robinhood trading as a proxy for attention-induced herding episodes and find that intense buying by Robinhood users forecasts negative returns with average 20-day abnormal returns being −4.7% for the top stocks purchased each day. Cookson et al. (2024) find similar evidence that higher attention on three major social media in the U.S. is associated with higher same-day returns, followed by a partial reversal over the next 10 to 20 days. Nevertheless, the return predictability on all stocks in this paper suggest that investors in the China and U.S. stock market pay attention

differently, as Da, Engelberg, and Gao (2011) find that retail attention-induced price pressure and reversal are only significant among smaller stocks and are largely absent for large stocks in the U.S. market. Adding to the literature, I document return predictability of unsophisticated retail attention among all stocks in the China stock market, which is dominated by retail investors.

[Insert Table II Here]

Figure 1 plots the coefficients $\beta 1_{t+j}$ of four attention measures in Table II. *Research Report Attention Index* and *News Attention Index* show similar pattern, both of which exhibit the most pronounced positive predictive power at T+1. *Community Attention Index* shows negative predictive power of return from T+1 onwards.

[Insert Figure 1 Here]

Figure 2 describes the cumulative effect of four attention measures in Table II. The cumulative effect at horizon k (x-axis) equals $\sum_{j=0}^{k} \beta 1_{t+j}$, where $\beta 1_{t+j}$ is the coefficient for day T+j from Table II. *Research Report Attention Index* and *News Attention Index* show positive cumulative effects, whereas *Research Report Attention Index* demonstrates larger magnitude than *News Attention Index*. *Announcement Attention Index* shows moderate positive cumulative effect. *Community Attention Index* shows negative cumulative effect.

[Insert Figure 2 Here]

ii. Buy-Sell Imbalance

Orders of small retail investors are in the wrong direction for future price movements. Jones et al. (2025) demonstrate that small retail order flows exhibit a momentum pattern—buying high and

selling low—effectively demanding liquidity and potentially leading to lower future returns. In contrast, large retail investors and institutional investors follow a contrarian pattern, i.e., buying low and selling high, providing liquidity to the market literally and is likely compensated positively. Clearly, Chinese retail investors have critical differences from U.S. markets. Unlike homogeneous retail investors in the U.S., small Chinese retail investors are momentum investors whose trades are negatively associated with future returns, while large Chinese retail investors are contrarian investors like institutional investors whose trades positively predict future returns.

Specifically, Jones et al. (2025) find that order imbalance is negatively related to future price movements for retail investors with accounts smaller than 3 million CNY, while order imbalance of institutional investors as well as retail investors higher than 3 million CNY accounts positively predict future return in China stock market. This suggests that higher ownership of smaller retail investors is related to overpricing in China stock markets. In contrast, U.S. retail investors are found to be contrarian investors, whose trades positively predict future returns (Kaniel, Saar and Titman, 2008; Kelley and Tetlock, 2013; Boehmer et al., 2021).

Table III focuses on the effect of attention on buy-sell imbalance of unsophisticated retail trades (small and middle trade), and sophisticated retail trades and institutional trades (large and extra-large trade). Specifically, I use transaction orders with fewer than 100,000 shares and a total value under CNY 200,000 as proxy for unsophisticated retail investors' trading activities. I use transaction orders with larger than or equal to 100,000 shares and a total value of CNY 200,000 as proxy for sophisticated retail and institutional investors' trading activities. The results reveal a striking contrast in how different attention translates into trading behavior over time.

Unsophisticated retail investors react positively to community attention (herding) but negatively to news and research reports initially. Sophisticated investors do the opposite — they buy on research and news but sell into community attention, consistent with liquidity provision or contrarian strategies.

Contemporaneously, the four attention dimensions show divergent effects. *Community Attention Index* is positively correlated with unsophisticated buy-sell imbalance but negatively correlated with sophisticated buy-sell imbalance on day T. In contrast, *Research Report, News, and Announcement Attention Indexes* are all negatively associated with unsophisticated BSI and positively associated with sophisticated BSI on day T. Specifically, a 1% increase in the *Community Attention Index* is associated with a 0.241 bps increase in unsophisticated BSI and a 0.605 bps decrease in sophisticated BSI on day T. In contrast, a 1% increase in *Research Report Attention Index* is associated with a 0.195 bps decrease in unsophisticated BSI and a 0.556 bps increase in sophisticated BSI on day T. *News Attention Index* shows a 0.249 bps decrease for retail and a 0.529 bps increase for sophisticated investors. *Announcement Attention Index* exhibits smaller but significant opposite-signed effects (negative for retail, positive for institutional). This is consistent with unsophisticated retail investors relying on noisy community sentiment, while sophisticated and institutional investors respond to fundamental information sources.

In terms of predictability, *Community Attention Index* positively predicts retail order imbalance for up to 10 trading days (T+1 through T+11, with peaks at T+1 and T+8–T+11, except T+4) and negatively predicts institutional order imbalance on T+1 before turning positive on T+2 through T+7. A 1% increase in Community Attention is associated with a 0.058 bps increase in

unsophisticated retail BSI and a 0.175 bps decrease in sophisticated BSI on T+1. This suggests that community-driven attention fuels buying pressure from unsophisticated retail investors, creating transitory overvaluation that sophisticated investors initially sell into. Interestingly, sophisticated investors begin buying from T+2 to T+7 when *Community Attention Index* rises, indicating they may later capitalize on retail-driven price movements.

In sharp contrast, a 1% increase in *Research Report Attention Index* predicts a 0.020 bps decrease in unsophisticated retail BSI and a 0.082 bps increase in sophisticated BSI on T+1. Sophisticated investors continue buying on T+2 before selling from T+4 to T+7. Unsophisticated retail investors begin buying only from T+3 onwards. This pattern suggests that institutional buying following research reports raises prices by a few basis points, which later attracts retail buyers, after which institutions sell to realize profits.

The results are consistent with the observation that trading is hazardous to most individual investors' wealth, while wealthy investors with portfolio values above the 99.5th percentile benefit from trading (Li et al., 2017). The wealthiest individuals often display hybrid strategies, behaving like institutions when buying (a sophisticated, information-driven act) but like the general retail population when selling (Ng and Wu, 2007). Ultimately, the key advantage for institutional and large retail investors is that they possess information advantages, either because of private access to valuable information or better skills in processing public information (Tan, Zhang and Zhang, 2024).

Furthermore, a 1% increase in News Attention Index predicts a statistically insignificant change in retail BSI on T+1, and a 0.033 bps increase in institutional BSI on T+1. Announcement

Attention predicts a −0.008 bps changes in retail BSI and a 0.018 bps increase in institutional BSI on T+1. These magnitudes are smaller than those of Community and Research Report attention but reinforce the same directional pattern. This is consistent with the trading behavior that unsophisticated retail investors rely on noisy information from community to trade while sophisticated retail and institutional investors obtain information from news, announcement and research reports. Sophisticated retail and institutional investors tend to trade in the opposite direction to unsophisticated retail investors.

[Insert Table III Here]

iii. Baidu Search Volume Index

Table IV controls for Baidu Search Volume Index (SVI) as an alternative measure of attention, and run horse-race regressions between attention indexes from Eastmoney and Baidu Search Volume Index. In Panel A, Baidu SVI is positively correlated with stock return, and predict stock return negatively up to 9 trading days. A one percentage increase in Baidu SVI is associated with a striking 180 bps increase in stock return contemporaneously, and cumulative 46.7 bps decrease from T+1 to T+9 trading days. This is consistent with Da et al. (2011) that an initial price increase is due to temporary price pressure, and it would revert in the long run. In Panel B, the effects of *Community Attention Index*, *News Attention Index*, *Announcement Attention Index*, and *Research Report Attention Index* also remain qualitatively similar with Table II after controlling for Baidu SVI. One exception is the flip of sign of *Community Attention Index* contemporaneously and for the next trading day (relative to Table II), as both Baidu Search Volume Index and *Community Attention Index* are populated by unsophisticated retail investors, therefore highly correlated.

Nevertheless, *Community Attention Index* still contains incremental information beyond Baidu SVI that it predicts return negatively from T+3 from T+10 whereas there is no such effect or weaker effect from Baidu SVI. The results suggest that attention from community, news, announcement and research reports in EastMoney App conveys incremental information compared with attention captured by Baidu Index. Therefore, this paper provides novel and direct measures of attention and suggests that investor attention can be measured more accurately and holistically beyond search engines such as Baidu.

[Insert Table IV Here]

IV. Attention and Sentiment

Loughran and Mcdonald (2011) develop positive and negative word lists based on 10-K sample in the U.S. and link word lists to 10-K filing returns and trading volume. Jiang, Meng and Tang (2021) construct a Chinese financial sentiment word dictionary based on Loughran and Mcdonald (2011) dictionary using word2vec algorithm and hand collection method, which positively predicts daily stock return. Tetlock (2007) measures media pessimism based on daily content from a popular Wall Street Journal column, and finds that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals. The results are consistent with theories of noise traders (De Long et al., 1990a) and liquidity traders (Campbell et al., 1993). Specifically, when noise traders experience a negative belief shock, they sell stocks to arbitrageurs, increasing volume and temporarily depressing returns. Returns rebound next period when there is a new belief shock. Da, Engelberg and Gao (2014) construct a measure of investor sentiment FEARS by

aggregating the volume of queries related to household concerns, which predicts short-term return reversals. Cookson et al. (2024) find that over two-thirds of the firm-day attention signal is common across the major social platforms (Twitter, StockTwits, and Seeking Alpha), while the common component of sentiment is weak.

Table V studies the interaction effect of sentiment Pos_Neg_{it} and various attention measures on stock return. Panel A studies the independent effect of sentiment. A one unit increase in sentiment is associated with 77.6 bps increase in return contemporaneously, and 3.9 bps increase next trading day. Panel B focuses on the interaction effect of *Community Attention Index*, *News Attention Index*, *Announcement Attention Index*, *Research Report Attention Index* with sentiment. The dependent variable is the return on day *t* for stock *i*, where k = 0, 1, 2…10.

$$
\begin{aligned}
Ret_{i,t+k} = a0 \quad & + \beta 1_{i,t} Attention_{i,t} + \beta 2_{i,t} Attention_{i,t} * High\ Sentiment_{i,t} \\
& + \beta 3_{i,t} High\ Sentiment_{i,t} \\
& + \beta 4_{i,t} \log_Size_{i,t} + \beta 5_{i,t} EP_ratio_{i,t} + \beta 6_{i,t} Turnover_{i,t} + \beta 7_{i,t} Ret_lag1_{i,t-1} \\
& + \beta 8_{i,t} Ret_lag2_{i,t-2} + \beta 9_{i,t} Ret_lag3_{i,t-3} + \beta 10_{i,t} Ret_lag4_{i,t-4} \\
& + \beta 11_{i,t} Ret_lag5_{i,t-5} + \varepsilon_{i,t+k}
\end{aligned}
$$

High sentiment universally amplifies next-day returns across all attention types on T+1 , with amplification effects ranging from 0.4 bps (*Announcement Attention Index*) to 7.3 bps (*Community Attention Index*). Different investor attention types show distinct return relationships: *Research Report Attention Index* provides the most persistent positive predictions while *Community Attention Index* exhibits sentiment-dependent reversals. This suggests that higher-quality attention

sources show greater effect persistence, suggesting investor clientele differences between speculative community and fundamental research analysis.

The economic significance of these findings is substantial, particularly for quantitative trading strategies and risk management. During high sentiment periods, a 1% increase in *Community Attention Index* generates a 71.3 bps return difference between high and low sentiment contemporaneously and 7.3 bps on T+1. The T+2 prediction result of *Community Attention Index* reveals a rapid decay in attention-driven returns. *Community Attention Index* maintains the T+2 predictive power at -5.9 bps during low sentiment periods and –5.2 bps during high sentiment periods, which exhibited sentiment-dependent reversals (positive at T during high sentiment, negative thereafter), reverts to an insignificant amplification effect on the T+2 trading day.

The sentiment-dependence in community attention further highlight significant mispricing during emotional market periods, creating both substantial risks for uninformed investors and profitable opportunities for sentiment-aware strategies. During periods of market euphoria, high noise attention coincides with buying pressure and rising prices. This is consistent with attention-induced herding. In a bullish environment, retail investors are more likely to interpret high community activity (posts, comments) as a positive signal, leading to coordinated buying. They are "chasing the buzz" and fueling momentum. However, this positive effect is transient and subsequently reverse on the next trading day, representing a short-term overreaction to non-fundamental, salience-based signals. In a bearish environment, community forums likely become echo chambers of concern, panic, or disappointment. Elevated attention reflects investors rushing to discuss losses, exit positions, or warn others, creating synchronized selling pressure. This

attention is driven by salience of negative stimuli. This suggests attention is now focused on selling, fear, or negative news. Noise attention from unsophisticated retail investors is reactive and sentiment-contingent. Its relationship with returns is not inherently positive or negative but depends on the prevailing market mood. This indicates that unsophisticated retail investors, as a group, do not use attention to uncover fundamental value but rather as a trigger for sentiment-driven herd behavior, buying in euphoria and selling in panic.

In sharp contrast, 1% increase in *Research Report Attention Index* produces 3.6 bps increase in returns the next trading day during low sentiment period and 5.7 bps during high sentiment period. *Research Report Attention Index*, proxy for sophisticated retail and institutional attention, positively predicts return on both high and low sentiment days. Regardless of the market mood, institutional attention predicts positive returns, providing evidence on fundamental information processing. Institutions pay attention to research reports to analyze a company's intrinsic value. If a report contains positive fundamental news (e.g., upgraded earnings forecast, strong operational metrics), institutions will buy, pushing the price up. This occurs irrespective of whether the broader market is bullish or bearish because the signal is stock-specific and fundamental. In a low-sentiment environment, this might represent a contrarian "buying the dip" opportunity based on their analysis. Institutional attention is proactive and information-based. Its consistent positive relationship with returns indicates that it is directed towards price-relevant fundamental information.

This result reveals a fundamental dichotomy in how different investor types process information and act upon it, depending on the market's emotional context. The sentiment regime

acts as a switch that alters the very nature of unsophisticated retail attention, while sophisticated retail and institutional attention remains consistently informed. The "dumb money" from the "smart money" is disentangled not by their identity, but by the channel of their attention. The opposing results also provide the micro-foundation for why average retail investors underperform and institutions outperform. Unsophisticated retailers are distracted by the sentiment-driven "noise" of the community, while institutions focus on the "signal" in research reports.

The persistence of *Research Report Attention Index* effects suggests sustainable alpha generation during both high and low sentiment period, whereas the quick reversal of *Community Attention Index* on the next trading day for high sentiment period indicates potential arbitrage opportunities through contrarian strategies. During low sentiment days, a 1% increase of Research Report Attention Index predicts cumulative 7.5 bps increase in returns over the next 1–10 trading days. During high sentiment days, a 1% increase of Research Report Attention Index predicts cumulative 9.4 bps increase in returns over the next 1–10 trading days. During low sentiment days, a 1% increase of Community Attention Index predicts cumulative -31.4 bps in returns over the next 1–10 trading days. In contrast, during high sentiment days, a 1% increase of Community Attention Index predicts cumulative -22.8 bps in returns over the same period. The exceptional persistence suggests that institutional investors and quantitative funds could potentially scale strategies around this signal.

News Attention Index shows consistent positive predictive effect for the next two trading days, which subsequently partially reverses for both high and low sentiment periods from T+3 onwards. *News Attention Index* shows a significant positive predictive effect only for the next trading day

(T+1) during high sentiment periods (4.2 bps), with an insignificant baseline effect during low sentiment (0.2 bps). By T+2, the effect turns negative (-1.0 bps for low sentiment, -0.4 bps for high sentiment), and this partial reversal continues from T+3 onwards. The pattern for news attention—immediate positive reaction, positive continuation the next day, followed by a reversal—is a short-lived informational cascade. On day T, positive news is released, which grabs investor attention, triggering immediate buying from investors who rush to capitalize on the new information (11.5 bps for low sentiment) and 23.5 bps for high sentiment). This is the initial "attention shock" amplified by high sentiment. The market incorporates the news and the associated optimism into the price very quickly. The price continues to rise on the following day. This is the informational cascade. The initial price increase itself acts as a new signal, attracting a second wave of investors who fear missing out (FOMO). This is often driven by retail investors or trend-following algorithms that react to the price momentum, not the original news itself. The news story may also be picked up and amplified by financial media throughout the day, sustaining attention. Two trading days afterwards, the market realizes that the initial news did not justify the entire price move. The buying pressure was largely "hot money"—non-fundamental, attention-driven flow that is quick to enter and quick to exit. As the news cycle moves on, attention fades, and this temporary demand evaporates, causing the price to fall back to its fundamental level. This is a clear overreaction followed by partial correction.

While news is a formal information source, its impact is also influenced by behavioral biases (like herding and overreaction) rather than pure fundamental analysis. It contains a significant "noise" component. This pattern provides trading opportunity for short-term traders and a trap for

slow-moving retail investors. It creates a predictable window of momentum (Day T to Day T+1) followed by a predictable reversal, allowing sophisticated players to "ride the wave" and then exit before the crash. Sentiment further amplifies the positive effects of news on Day T+1 and T+2 by 4 bps and 0.6 bps. This finding powerfully demonstrates that even attention to credible information sources can generate temporary mispricing if the information is easily digestible and widely disseminated, making it prone to be traded on by a large, non-professional investors.

Announcement Attention Index shows weaker effects, with a minimal 0.4 bps sentiment amplification on the next trading day. The pattern for *Announcement Attention Index* indicates that initial reactions are swayed by sentiment, but the market ultimately corrects to properly value the fundamental information contained in corporate announcements. In a bullish market, high attention to corporate announcements is associated with immediate price increases, as the market is predisposed to interpret news—even neutral or mildly positive announcements—in a positive light. This is consistent with sentiment-amplified reaction. The combination of a concrete information release (the announcement) and high sentiment lead to a strong, immediate bullish reaction. During low sentiment period, high attention to announcements is associated with delayed reactions at T+3 and T+4, possibly due to delayed information acquisition. The delayed effects of *Announcement Attention Index* are consistent with post-earnings announcement drift.

Corporate announcements are a direct source of fundamental information, but they are consumed by a heterogeneous audience. The initial reaction is a mix of sophisticated and unsophisticated trading, with sentiment coloring the interpretation. In low sentiment periods, the initial market reaction is inefficient, with the market being too pessimistic. This pattern is

consistent with underreaction followed by delayed correction. If the corporate announcements contained fundamentally positive or neutral information that was initially overlooked or misinterpreted due to the gloomy market mood, sophisticated investors would begin to buy. Their buying over the subsequent days corrects the initial mispricing, leading to the positive predictive return. This creates a clear arbitrage opportunity. The delayed effect after three days on low-sentiment periods is a pattern that can be exploited by investors who can correctly assess the fundamental value of the announcement and withstand the short-term sentiment pressure. In both sentiment regimes, the initial market reaction is inefficient. The direction of this correction is always positive, implying that the announcements that attract high attention are, on average, conveying positive fundamental news. This reveals a profound and exploitable market anomaly. It shows that while sentiment creates noise and temporary mispricing around corporate announcements, the fundamental signal is so powerful that it forces a predictable correction.

In all panels, Pos_Neg_{it} predicts return positively at T and T+1, while the effect is insignificant at longer horizons, suggesting the informativeness of sentiment as it might contain return-relevant information in the short run. The result is consistent with Cookson et al. (2024) that social media sentiment predicts positive next-day returns. Overall, the market impact of attention depends on the prevailing sentiment environment. The findings elegantly bridge the literature on investor sentiment, limited attention, and market efficiency.

[Insert Table V Here]

Figures 3 a-d plot the coefficients of various attention measures in the high sentiment and low sentiment days respectively. In Figure 3a, *Community Attention Index* shows different pattern

under high sentiment and low sentiment period. During high sentiment period, Community Attention Index are positively correlated with stock return contemporaneously which subsequently turns negative (reversal). During low sentiment period, *Community Attention Index* negatively predicts stock return that subsequently diminishes. This suggests that retail investors who tend to be noise traders pay attention differently under high and low sentiment periods. Figure 3b shows that *Research Report Attention Index* consistently positively predict return, and the positive effects are enhanced during high sentiment periods without reversal. This suggests that sophisticated retail investors and institutional investors trade on information. Figure 3c shows that *News Attention Index* exhibits partial correction. Figure 3d indicates delayed effects from *Announcement Attention Index.*

[Insert Figure 3 Here]

Figures 4 a-d plot the cumulative effects of various attention measures in the high sentiment and low sentiment days respectively. *Research Report Attention Index*, *News Attention Index* and *Announcement Attention Index* show positive cumulative effects under both high and low sentiment periods in Figure 4b, 4c and 4d, suggesting the information contents from these three sources. In sharp contrast, *Community Attention Index s*hows contrasting pattern of positive cumulative effects under high sentiment periods and negative cumulative effects under low sentiment period in Figure 4a, suggesting that unsophisticated investors are highly influenced by sentiment.

[Insert Figure 4 Here]

V. Conclusion

A central challenge in attention literature is to identify the exact measures of attention due to its inherent unobservability. This paper tackles this issue by leveraging a novel, high-frequency dataset from EastMoney, one of China's largest retail trading platforms, to introduce direct, behavior-based measures of attention across distinct informational channels. Depending on where attention is focused on, this paper studies the effect of community, research reports, news and announcement attention on stock return. It is found that community attention negatively predicts return, reflecting retail noise. In contrast, research report attention positively predict return without reversal, signaling institutional incorporation of fundamental information. News attention positively predicts return but subsequently reverses, suggesting the potential overreaction towards news. Announcement attention exhibits a delayed positive effect, suggesting initial underreaction that is later corrected. All four attention measures capture additional information beyond current Baidu Search Volume Index.

In addition to separate the attention from the sources, this paper proposes a direct measure of unsophisticated retail versus sophisticated retail and institutional investor attention in the context of Chinese stock market: unsophisticated retail attention (unsophisticated attention) is proxied by *Community Attention Index* while sophisticated retail and institutional attention (sophisticated attention) is measured by *Research Report Attention Index*. Unsophisticated retail investors are more likely to obtain information from community (investor forums), while sophisticated retail and institutional investors are more likely to read research reports. Noise trades from unsophisticated retail investors usually predicts return negatively while sophisticated trades from

sophisticated retail and institutional investors predict future return positively. The results in the paper also suggest that research reports contain information that is beyond news and announcements. The divergent return predictability of these measures directly mirrors the documented sophistication gap in the Chinese market. Unsophisticated retail investors, drawn to the community, generate negative pressure; institutions, focused on research reports, drive positive, permanent price changes.

The interaction analysis between sentiment and attention reveals the dual nature of attention: attention works differently depending on market sentiment. It suggests that attention measures are not uniformly "good" or "bad". The context (sentiment) also determines the direction of attention effects. Specifically, *Community Attention Index s*hows contrasting pattern of positive cumulative effects under high sentiment periods and negative cumulative effects under low sentiment period, suggesting that unsophisticated investors are highly influenced by sentiment.

If average retail attention were directed towards research reports rather than community sentiment, market efficiency could be significantly enhanced. This paper not only provides a new framework for measuring attention but also identifies a viable pathway for improving price discovery in emerging markets. Further research might delve more into the attention spillover between noise traders and sophisticated traders.

References:

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The journal of finance*, *59*(3), 1259-1294.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *The journal of finance*, *77*(6), 3141-3190.

Barber, B. M., & Odean, T. (2002). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of finance*, *55*(2), 773-806.

Barber, B. M., & Odean, T. (2001). The internet and the investor. *Journal of Economic Perspectives*, *15*(1), 41-54.

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

Barber, B. M., Odean, T., & Zhu, N. (2009). Systematic noise. *Journal of Financial Markets*, *12*(4), 547-569.

Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review*, *93*(3), 25-57.

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of financial studies*, *30*(9), 3009-3047.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Tracking retail investor activity. *The journal of finance*, *76*(5), 2249-2305.

Booker, A., Curtis, A., & Richardson, V. J. (2023). Investor disagreement, disclosure processing costs, and trading volume evidence from social media. *The Accounting Review*, *98*(1), 109-137.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The quarterly journal of economics*, *127*(3), 1243-1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and asset prices. *American Economic Review*, *103*(3), 623-628.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803-843.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2020). Memory, attention, and choice. *The quarterly journal of economics*, *135*(3), 1399-1442.

Bui, D. G., Huang, S., Lin, C.-Y., & Lin, T.-C. (2021). Tradin'in the Rain: Attention Allocation and Investment Performance. *Available at SSRN 3881853*.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *The quarterly journal of economics*, *108*(4), 905-939.

Chang, Y.-C., Shao, R., & Wang, N. (2022). Can stock message board sentiment predict future returns? Local versus nonlocal posts. *Journal of Behavioral and Experimental Finance*, *34*, 100625.

Chemmanur, T., & Yan, A. (2009). Product market advertising and new equity issues. *Journal of financial economics*, *92*(1), 40-65.

Chen, H., De, P., Hu, Y., & Hwang, B.-H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of financial studies*, *27*(5), 1367-1403.

Chen, J., Tang, G., Yao, J., & Zhou, G. (2022). Investor attention and stock returns. *Journal of Financial and Quantitative Analysis*, *57*(2), 455-484.

Chen, X., An, L., Wang, Z., & Yu, J. (2023). Attention spillover in asset pricing. *The journal of finance*, *78*(6), 3515-3559.

Chordia, T., & Swaminathan, B. (2000). Trading volume and cross-autocorrelations in stock returns. *The journal of finance*, *55*(2), 913-935.

Cookson, J. A., Lu, R., Mullins, W., & Niessner, M. (2024). The social signal. *Journal of Financial Economics*, 158, 103870.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The journal of finance*, *66*(5), 1461-1499.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *The Review of financial studies*, *28*(1), 1-32.

Da, Z., Hua, J., Hung, T. C.-C., & Peng, L. (2025). Market returns and a tale of two types of attention. Management Science, 71(12), 10505–10537.

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*(9), 1375-1388.

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

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990b). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. The Journal of Finance 45: 379-95.

Dorn, A. J., Dorn, D., & Sengmueller, P. (2015). Trading as gambling. *Management Science*, *61*(10), 2376-2393.

Eaton, G. W., Green, T. C., Roseman, B. S., & Wu, Y. (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of financial economics*, *146*(2), 502-528.

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

Focke, F., Ruenzi, S., & Ungeheuer, M. (2020). Advertising, attention, and financial markets. *The Review of Financial Studies*, 33(10), 4676–4720.

Gao, X., & Lin, T.-C. (2015). Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments. *The Review of financial studies*, *28*(7), 2128-2166.

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The high-volume return premium. *The journal of finance*, *56*(3), 877-919.

Gondhi, N. (2023). Rational inattention, misallocation, and the aggregate economy. *Journal of monetary Economics*, *136*, 50-75.

Griffin, J. M., Harris, J. H., & Topaloglu, S. (2003). The dynamics of institutional and individual trading. *The journal of finance*, *58*(6), 2285-2320.

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *The Review of financial studies*, *17*(2), 439-461.

Gu, C., & Kurov, A. (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking & Finance*, *121*, 105969.

Guo, R., Jia, D., & Sun, X. (2023). Information acquisition, uncertainty reduction, and pre-announcement premium in China. *Review of Finance*, *27*(3), 1077-1118.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The journal of finance*, *64*(5), 2289-2325.

Hirshleifer, D., & Sheng, J. (2022). Macro news and micro news: complements or substitutes? *Journal of financial economics*, *145*(3), 1006-1024.

Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics*, *36*(1-3), 337-386.

Hong, H., Jiang, W., Wang, N., & Zhao, B. (2014). Trading for status. *The Review of financial studies*, *27*(11), 3171-3212.

Hou, K., Peng, L., & Xiong, W. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. *Available at SSRN 976394*.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of financial studies*, *28*(3), 791-837.

Huang, S., Huang, Y., & Lin, T.-C. (2019). Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of financial economics*, *132*(2), 369-383.

Huang, Y., Qiu, H., & Wu, Z. (2016). Local bias in investor attention: Evidence from China's Internet stock message boards. *Journal of Empirical finance*, *38*, 338-354.

Jiang, L., Liu, J., Peng, L., & Wang, B. (2022). Investor attention and asset pricing anomalies. *Review of Finance*, *26*(3), 563-593.

Jiang, L., Liu, J., & Yang, B. (2019). Communication and comovement: Evidence from online stock forums. *Financial Management*, *48*(3), 805-847.

Jiang, F., Meng, L., & Tang, G. (2021). Media text sentiment and stock return prediction. *China Economic Quarterly*, 21(4), 1323-1344.

Jones, C. M., Shi, D., Zhang, X., & Zhang, X. (2025). Retail trading and return predictability in China. *Journal of Financial and Quantitative Analysis*, *60*(1), 68-104.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The journal of finance*, *63*(1), 273-310.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? Insights from retail orders and stock returns. *The journal of finance*, *68*(3), 1229-1265.

Li, J., & Yu, J. (2012). Investor attention, psychological anchors, and stock return predictability. *Journal of financial economics*, *104*(2), 401-419.

Li, X., Geng, Z., Subrahmanyam, A., & Yu, H. (2017). Do wealthy investors have an informational advantage? Evidence based on account classifications of individual investors. *Journal of Empirical finance*, *44*, 1-18.

Li, X., Shen, D., & Zhang, W. (2018). Do Chinese internet stock message boards convey firm-specific information? *Pacific-Basin Finance Journal*, *49*, 1-14.

Liu, H., Peng, L., & Tang, Y. (2023). Retail attention, institutional attention. *Journal of Financial and Quantitative Analysis*, *58*(3), 1005-1038.

Lou, D. (2014). Attracting investor attention through advertising. *The Review of financial studies*, *27*(6), 1797-1829.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The journal of finance*, *66*(1), 35-65.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of economic theory*, *145*(5), 1837-1864.

Ng, L., & Wu, F. (2007). The trading behavior of institutions and individuals in Chinese equity markets. *Journal of Banking & Finance*, *31*(9), 2695-2710.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, *40*(2), 307-329.

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of financial economics*, *80*(3), 563-602.

Peress, J., & Schmidt, D. (2020). Glued to the TV: Distracted noise traders and stock market liquidity. *The journal of finance*, *75*(2), 1083-1133.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: what is the role of the United States? *The journal of finance*, *68*(4), 1633-1662.

Sabherwal, S., Sarkar, S. K., & Zhang, Y. (2011). Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news. *Journal of Business Finance & Accounting*, *38*(9-10), 1209-1237.

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical finance*, *14*(5), 590-610.

Shleifer, A., and L. H. Summers. 1990. The Noise Trader Approach to Finance. Journal of Economic Perspectives 4: 19-33.

Shleifer, A., and R. W. Vishny. 1997. The Limits of Arbitrage. The Journal of Finance 52: 35-55.

Sicherman, N., Loewenstein, G., Seppi, D. J., & Utkus, S. P. (2016). Financial attention. The *Review of Financial Studies*, 29(4), 863–897.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665-690.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, *96*(2), 158-163.

Song, C. (2020). Financial illiteracy and pension contributions: A field experiment on compound interest in China. *The Review of financial studies*, *33*(2), 916-949.

Tan, L., Zhang, X., & Zhang, X. (2024). Retail and Institutional Investor Trading Behaviors: Evidence from China. *Annual Review of Financial Economics*, 16, 459-483.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The journal of finance*, *62*(3), 1139-1168.

Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information? *The*

Review of financial studies, *24*(5), 1481-1512.
Titman, S., Wei, C., & Zhao, B. (2022). Corporate actions and the manipulation of retail investors in China: An analysis of stock splits. *Journal of financial economics*, *145*(3), 762-787.
Van Nieuwerburgh, S., & Veldkamp, L. (2010). Information acquisition and under-diversification. *The Review of Economic Studies*, *77*(2), 779-805.
Wang, K., Yu, X., & Zhang, B. (2023). Panda games: corporate disclosure in the eclipse of search. *Management Science*, *69*(6), 3263-3284.
Welch, I. (2022). The wisdom of the Robinhood crowd. *The journal of finance*, *77*(3), 1489-1527.
Xu, Y., Xuan, Y., & Zheng, G. (2021). Internet searching and stock price crash risk: Evidence from a quasi-natural experiment. *Journal of financial economics*, *141*(1), 255-275.
Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of financial economics*, *116*(3), 548-564.

Table I Summary Statistics

Panel A reports mean, median, minimum, maximum, standard deviation, and number of observations of main variables. Panel B reports pairwise correlation between the attention measures.

Variable	Mean	Median	Min	Max	SD	N
Community Attention Index	-0.058	-0.038	-10.737	10.253	0.935	2952905
Research Report Attention Index	-0.161	-0.015	-8.489	8.434	1.497	2952905
News Attention Index	0.249	-0.027	-7.267	14.523	1.414	2952905
Announcement Attention Index	0.253	-0.043	-8.639	9.475	1.519	2952905
Stock Return	0.032	0.000	-82.581	20.833	3.019	2952905
BSI_Amount_Unsophisticated	0.018	0.016	-1.000	1.000	0.071	2952905
BSI_Amount_Sophisticated	-0.067	-0.060	-1.000	1.000	0.235	2952905
Baidu SVI	0.120	0.075	-1.995	3.734	0.257	1801280
Sentiment	0.191	0.201	-1.000	1.000	0.528	2631137
EP_Ratio	0.087	0.026	0.000	409.836	3.713	2952905
Turnover	0.018	0.010	0.000	0.581	0.023	2952905
Log (Market Value in Thousands of CNY)	15.917	15.679	11.137	21.907	1.086	2952905

Panel B Correlation

	Community Attention Index	Research Report Attention Index	News Attention Index	Announcement Attention Index
Community Attention Index	1.0000			
Research Report Attention Index	0.2812	1.0000		
News Attention Index	0.3833	0.2838	1.0000	
Announcement Attention Index	0.3042	0.1876	0.4167	1.0000

Table II Attention and Stock Return

This table studies the effect of attention measures on stock return., including the effect of community, news, announcement and research attention on stock return. Community Attention Index is measured by number of posts and replies in a firm-specific stock message board on Guba Eastmoney. Research Report Attention Index is derived from number of reads and comments on professional research reports. News Attention Index is based on the number of news article reads and associated comments for a given stock. Announcement Attention Index is calculated from the number of reads and comments on corporate announcements. Control variables include five lags of return, logarithm of market size (Size), earnings to price ratio and turnover to control for effect of size, value, and liquidity on stock return, consistent with Jones et al. (2025). This table adopts Fama-MacBeth regression.

$$Ret_{i,t+k} = a0 \quad + \beta 1_{i,t} Attention_{i,t} + \beta 2_{i,t} \log_Size_{i,t} + \beta 3_{i,t} EP_ratio_{i,t} + \beta 4_{i,t} Turnover_{i,t} + \beta 5_{i,t} Ret_lag1_{i,t-1}$$
$$+ \beta 6_{i,t} Ret_lag2_{i,t-2} + \beta 7_{i,t} Ret_lag3_{i,t-3} + \beta 8_{i,t} Ret_lag4_{i,t-4} + \beta 9_{i,t} Ret_lag5_{i,t-5} + \varepsilon_{i,t+k}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Return											
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10	T+11
Community Attention Index	0.030***	-0.046***	-0.058***	-0.049***	-0.023***	-0.020***	-0.022***	-0.011**	-0.012***	-0.016***	-0.012***	-0.005
	(4.00)	(-4.59)	(-11.39)	(-10.24)	(-5.16)	(-4.61)	(-5.10)	(-2.55)	(-2.89)	(-3.91)	(-2.97)	(-1.31)
Research Report Attention Index	0.137***	0.213***	0.044***	0.011***	0.008***	0.005***	0.000	0.002	0.004**	0.004**	0.003	0.000
	(55.21)	(66.63)	(21.52)	(6.22)	(4.32)	(2.81)	(0.18)	(1.17)	(2.34)	(2.02)	(1.63)	(0.19)
News Attention Index	0.140***	0.184***	0.021***	-0.008***	-0.008***	-0.005***	-0.007***	-0.003**	-0.001	-0.002	-0.001	-0.004***
	(57.29)	(67.77)	(11.24)	(-4.47)	(-4.46)	(-3.14)	(-4.52)	(-2.14)	(-0.95)	(-1.10)	(-0.71)	(-2.58)
Announcement Attention Index	0.012***	0.006***	0.001	0.002	0.004***	0.003*	0.002	-0.000	0.001	0.002*	0.002	0.001
	(8.36)	(4.17)	(0.72)	(1.16)	(3.21)	(1.92)	(1.32)	(-0.02)	(0.39)	(1.70)	(1.31)	(0.92)
Return_lag1	-0.023***	0.010***	-0.000	-0.000	-0.000	-0.009***	0.005**	0.002	-0.002	0.002	-0.004*	0.002
	(-8.00)	(3.36)	(-0.04)	(-0.04)	(-0.06)	(-3.93)	(2.08)	(1.03)	(-1.13)	(0.73)	(-1.92)	(1.09)
Return_lag2	-0.038***	-0.007***	0.001	0.002	-0.008***	0.006***	0.004*	-0.001	0.003	-0.003	0.002	0.001
	(-14.70)	(-2.61)	(0.41)	(0.84)	(-3.40)	(2.72)	(1.69)	(-0.57)	(1.65)	(-1.21)	(1.03)	(0.62)
Return_lag3	-0.034***	-0.008***	0.001	-0.008***	0.006**	0.004*	-0.001	0.003	-0.003	0.002	0.001	-0.002
	(-14.02)	(-3.44)	(0.53)	(-3.55)	(2.50)	(1.72)	(-0.52)	(1.61)	(-1.19)	(0.81)	(0.63)	(-0.95)
Return_lag4	-0.029***	-0.007***	-0.009***	0.006**	0.003	-0.001	0.003	-0.002	0.001	0.001	-0.002	-0.001
	(-12.56)	(-2.97)	(-3.95)	(2.51)	(1.24)	(-0.34)	(1.60)	(-1.00)	(0.65)	(0.42)	(-0.96)	(-0.52)
Return_lag5	-0.034***	-0.014***	0.005**	0.003	-0.001	0.003	-0.001	0.002	0.001	-0.002	-0.001	-0.002
	(-15.82)	(-6.61)	(2.42)	(1.33)	(-0.46)	(1.59)	(-0.63)	(0.94)	(0.56)	(-1.25)	(-0.45)	(-0.93)
EP_Ratio_lag	0.400***	-0.127	-0.054	-0.019	-0.052	-0.070	-0.035	-0.059	-0.061	-0.056	-0.054	-0.035
	(3.81)	(-1.20)	(-0.50)	(-0.17)	(-0.48)	(-0.65)	(-0.33)	(-0.55)	(-0.57)	(-0.52)	(-0.50)	(-0.32)
Turnover_lag	17.544***	-8.349***	-4.103***	-2.436***	-2.525***	-2.465***	-1.854***	-1.876***	-1.869***	-1.787***	-1.569***	-1.426***
	(37.96)	(-19.33)	(-10.16)	(-6.24)	(-6.51)	(-6.53)	(-4.89)	(-5.01)	(-5.07)	(-4.87)	(-4.35)	(-3.99)
Log_Marketvalue_lag	0.081***	-0.071***	-0.032***	-0.021**	-0.022***	-0.020**	-0.017**	-0.019**	-0.020**	-0.019**	-0.019**	-0.017*
	(10.27)	(-8.77)	(-3.86)	(-2.49)	(-2.60)	(-2.32)	(-2.02)	(-2.22)	(-2.26)	(-2.18)	(-2.09)	(-1.91)
Constant	-1.634***	1.214***	0.553***	0.372**	0.406***	0.378**	0.324**	0.358**	0.371**	0.361**	0.347**	0.328**
	(-11.79)	(8.41)	(3.83)	(2.53)	(2.71)	(2.49)	(2.13)	(2.34)	(2.38)	(2.31)	(2.21)	(2.06)
Observations	2,952,905	2,952,990	2,950,028	2,947,055	2,944,069	2,941,077	2,938,091	2,935,112	2,932,135	2,929,162	2,926,192	2,923,232

R-squared	0.134	0.122	0.086	0.082	0.078	0.076	0.074	0.073	0.071	0.070	0.069	0.069
Number of groups	1,395	1,395	1,394	1,393	1,392	1,391	1,390	1,389	1,388	1,387	1,386	1,385

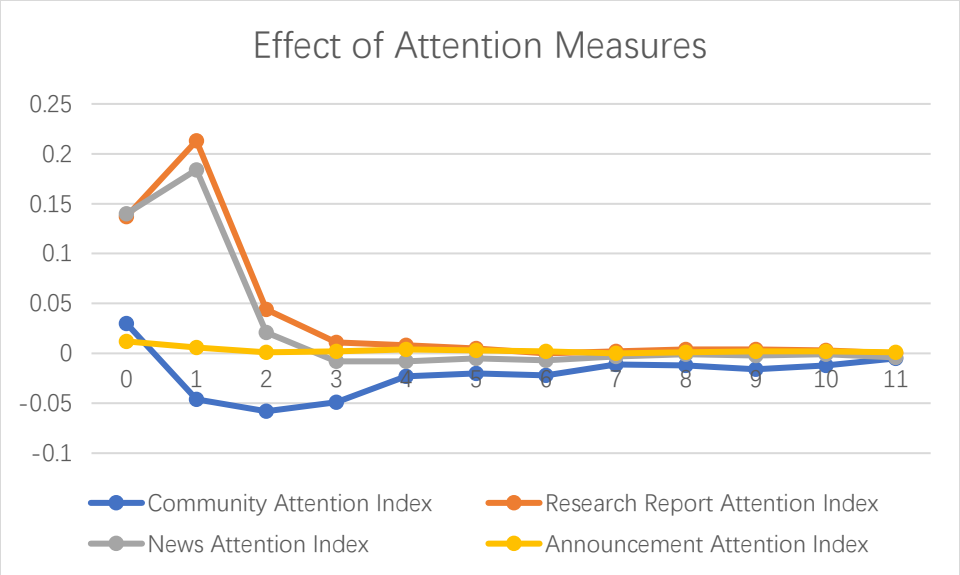


Figure 1 Coefficients of Attention Indexes on Stock Returns

This figure plots the Fama-MacBeth regression coefficients $\beta 1_{t+j}$ from Table II, showing the contemporaneous and predictive effects of Community Attention Index, Research Report Attention Index, News Attention Index, and Announcement Attention Index on stock returns from day T (contemporaneous) through T+11. Positive coefficients indicate positive return predictability, while negative coefficients indicate negative return predictability.

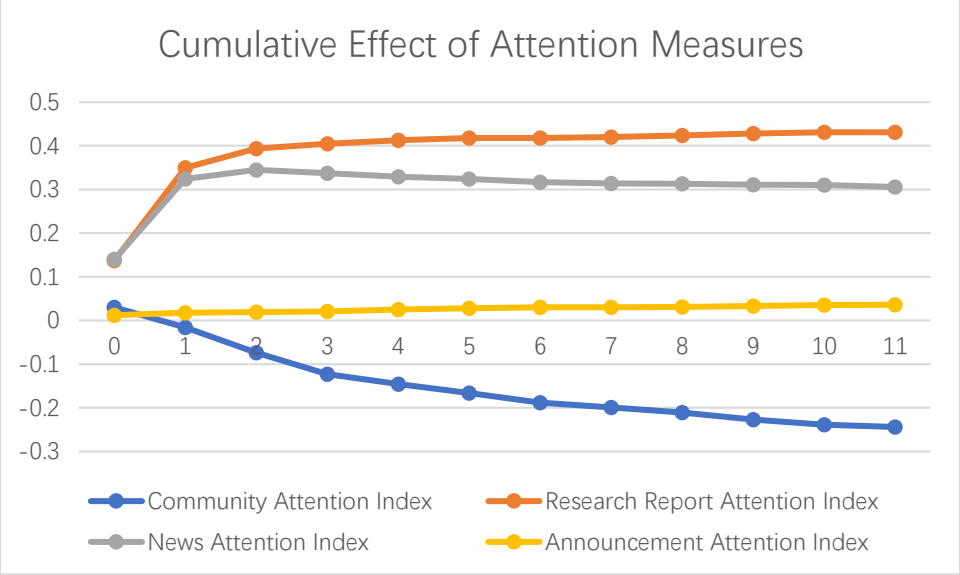


Figure 2 Cumulative Coefficients of Attention Indexes on Stock Returns

For each attention measure, cumulative coefficients of Community Attention Index, Research Report Attention Index, News Attention Index, and Announcement Attention Index are computed separately as the sum of individual Fama-MacBeth regression coefficients from day T through each subsequent horizon. Specifically, the cumulative effect at horizon k (x-axis) equals $\sum_{j=0}^{k} \beta 1_{t+j}$, where $\beta 1_{t+j}$ is the coefficient for day T+j from Table II. This summation captures the total return effect from a one-unit increase in the attention measure over a k-day holding period, allowing assessment of both the magnitude and persistence (or reversal) of return predictability. For example, a cumulative coefficient that increases monotonically indicates sustained positive effects, while a cumulative coefficient that rises then falls indicates temporary price pressure followed by reversal.

Table III Attention and Buy-Sell Imbalance

This table studies the effect of attentions on buy-sell imbalance of unsophisticated retail as well as sophisticated retail and institutional investors respectively. Panel A focuses on unsophisticated retail investors while Panel B focuses on sophisticated retail and institutional investors. I follow Kumar and Lee (2006) to construct Buy-Sell Imbalance (BSI). The trading activities information is obtained from EastMoney moneyflow. I use transaction orders with fewer than 100,000 shares and a total value under CNY 200,000 as proxy for unsophisticated retail investors' trading activities. I use transaction orders with larger than or equal to 100,000 shares and a total value of CNY 200,000 as proxy for sophisticated retail and institutional investors' trading activities. This is consistent with the classification in EastMoney that small trade and middle trade are proxy for small retail trading, while large trade and extra-large trade are proxy for large retail and institutional trading. Control variables include five lags of buy-sell imbalance, logarithm of market size (Size), earnings to price ratio and Turnover as control variables to control for effect of size, value, and liquidity on stock return. This table adopts OLS regression controlling for stock and date fixed effects, with standard error clustered at date level. BSI is constructed based on amount of transaction are constructed as follows and reported in percentage in the regressions:

$$BSI_Amount_{it} = \frac{Buy_Amount_{it} - Sell_Amount_{it}}{Buy_Amount_{it} + Sell_Amount_{it}}$$

Panel A Unsophisticated Retail Buy-Sell Imbalance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	BSI_Amount_Unsophisticated Unsophisticated Retail Investors											
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10	T+11
Community Attention Index	0.241***	0.058***	0.000	-0.006	-0.009*	-0.001	-0.001	0.006	0.018***	0.023***	0.012**	0.021***
	(48.52)	(11.62)	(0.05)	(-1.29)	(-1.81)	(-0.18)	(-0.24)	(1.12)	(3.49)	(4.51)	(2.30)	(4.13)
Research Report Attention Index	-0.195***	-0.020***	-0.003	0.006**	0.014***	0.022***	0.020***	0.017***	0.012***	0.012***	0.012***	0.006*
	(-66.53)	(-6.94)	(-1.11)	(2.11)	(4.61)	(7.46)	(6.56)	(5.62)	(4.09)	(3.89)	(3.98)	(1.86)
News Attention Index	-0.249***	0.003	0.008**	0.001	0.004	0.005	0.003	0.000	-0.006*	-0.009***	0.000	0.000
	(-75.39)	(0.82)	(2.25)	(0.43)	(1.10)	(1.56)	(0.98)	(0.01)	(-1.85)	(-2.67)	(0.01)	(0.08)
Announcement Attention Index	-0.018***	-0.008***	-0.002	-0.009***	-0.005*	-0.004	-0.002	-0.003	0.001	0.001	0.000	-0.001
	(-6.12)	(-2.66)	(-0.65)	(-3.07)	(-1.69)	(-1.35)	(-0.79)	(-1.01)	(0.35)	(0.29)	(0.11)	(-0.23)
BSI_Amount_lag1	0.139***	0.140***	0.081***	0.054***	0.038***	0.022***	0.019***	0.021***	0.018***	0.018***	0.016***	0.015***
	(238.10)	(241.05)	(138.16)	(91.11)	(63.69)	(36.74)	(32.51)	(36.05)	(30.35)	(31.12)	(27.57)	(25.59)
BSI_Amount_lag2	0.061***	0.067***	0.046***	0.033***	0.019***	0.019***	0.022***	0.018***	0.019***	0.016***	0.015***	0.014***
	(104.65)	(114.36)	(77.81)	(55.86)	(32.75)	(32.71)	(36.24)	(29.69)	(31.34)	(27.01)	(25.27)	(24.14)
BSI_Amount_lag3	0.035***	0.039***	0.029***	0.016***	0.017***	0.021***	0.017***	0.018***	0.015***	0.014***	0.013***	0.014***
	(59.13)	(66.37)	(48.99)	(27.84)	(29.51)	(35.21)	(28.16)	(29.75)	(25.33)	(23.02)	(22.06)	(24.02)
BSI_Amount_lag4	0.022***	0.026***	0.015***	0.017***	0.021***	0.017***	0.018***	0.015***	0.014***	0.013***	0.014***	0.010***
	(38.60)	(44.88)	(26.00)	(28.01)	(34.99)	(29.08)	(30.21)	(25.62)	(23.42)	(21.81)	(24.01)	(16.40)
BSI_Amount_lag5	0.012***	0.016***	0.018***	0.022***	0.019***	0.020***	0.017***	0.015***	0.015***	0.015***	0.011***	0.011***
	(21.60)	(27.28)	(30.92)	(37.67)	(32.55)	(34.51)	(29.06)	(26.33)	(25.02)	(25.99)	(18.33)	(19.29)
EP_Ratio_lag	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.66)	(-0.43)	(-0.51)	(-0.50)	(-0.55)	(-0.49)	(-0.63)	(-0.45)	(-0.51)	(-0.48)	(-0.44)	(-0.43)
Turnover_lag	3.885***	15.817***	8.863***	8.262***	8.987***	9.050***	8.095***	8.218***	7.918***	7.821***	7.404***	7.139***
	(18.09)	(73.68)	(40.88)	(38.00)	(41.30)	(41.56)	(37.15)	(37.70)	(36.30)	(35.84)	(33.91)	(32.68)
Log_Marketvalue_lag	-0.151***	0.565***	0.727***	0.756***	0.751***	0.739***	0.714***	0.684***	0.674***	0.658***	0.652***	0.644***
	(-11.00)	(41.18)	(52.39)	(54.32)	(53.86)	(52.94)	(51.15)	(48.97)	(48.20)	(47.00)	(46.55)	(45.96)
Constant	3.695***	-8.000***	-10.272***	-10.638***	-10.515***	-10.296***	-9.875***	-9.392***	-9.210***	-8.942***	-8.836***	-8.705***
	(16.96)	(-36.70)	(-46.65)	(-48.15)	(-47.53)	(-46.49)	(-44.56)	(-42.36)	(-41.50)	(-40.26)	(-39.74)	(-39.10)

Observations	2,954,180	2,951,303	2,948,330	2,945,344	2,942,352	2,939,366	2,936,387	2,933,410	2,930,437	2,927,467	2,924,507	2,921,550
R-squared	0.091	0.090	0.070	0.065	0.062	0.061	0.061	0.060	0.060	0.060	0.059	0.059
Stock FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Panel B Sophisticated Retail and Institutional Buy-Sell Imbalance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	BSI_Amount_Sophisticated											
	Sophisticated Retail and Institutional Investors											
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10	T+11
Community Attention Index	-0.605***	-0.175***	0.058***	0.098***	0.094***	0.082***	0.069***	0.032*	-0.021	-0.064***	-0.045***	-0.065***
	(-36.11)	(-10.40)	(3.42)	(5.75)	(5.49)	(4.80)	(4.06)	(1.88)	(-1.22)	(-3.74)	(-2.60)	(-3.76)
Research Report Attention Index	0.556***	0.082***	0.037***	0.014	-0.021**	-0.057***	-0.040***	-0.028***	-0.011	-0.008	-0.005	0.006
	(56.08)	(8.27)	(3.67)	(1.42)	(-2.07)	(-5.68)	(-3.97)	(-2.76)	(-1.07)	(-0.78)	(-0.51)	(0.55)
News Attention Index	0.529***	0.033***	-0.005	0.001	0.005	-0.009	-0.002	-0.008	0.014	0.010	-0.012	-0.002
	(47.66)	(2.97)	(-0.44)	(0.12)	(0.47)	(-0.77)	(-0.18)	(-0.69)	(1.25)	(0.86)	(-1.03)	(-0.22)
Announcement Attention Index	0.042***	0.018*	-0.002	0.028***	0.019*	0.026***	0.001	0.007	-0.007	0.012	0.004	-0.007
	(4.28)	(1.82)	(-0.17)	(2.84)	(1.95)	(2.60)	(0.09)	(0.71)	(-0.70)	(1.14)	(0.37)	(-0.66)
BSI_Amount_lag1	0.079***	0.085***	0.049***	0.036***	0.026***	0.016***	0.017***	0.018***	0.016***	0.017***	0.015***	0.015***
	(134.35)	(144.19)	(83.75)	(61.53)	(44.60)	(27.34)	(27.95)	(30.14)	(26.46)	(29.30)	(25.01)	(24.72)
BSI_Amount_lag2	0.039***	0.043***	0.032***	0.023***	0.014***	0.016***	0.017***	0.014***	0.016***	0.014***	0.014***	0.014***
	(66.24)	(73.62)	(54.34)	(38.76)	(23.49)	(26.20)	(28.53)	(24.20)	(27.13)	(22.82)	(23.19)	(22.49)
BSI_Amount_lag3	0.026***	0.030***	0.021***	0.013***	0.015***	0.017***	0.014***	0.016***	0.013***	0.013***	0.013***	0.011***
	(44.51)	(50.31)	(36.24)	(22.13)	(25.00)	(28.27)	(23.53)	(26.77)	(22.37)	(22.34)	(21.99)	(18.19)
BSI_Amount_lag4	0.017***	0.020***	0.013***	0.014***	0.017***	0.014***	0.016***	0.014***	0.013***	0.013***	0.011***	0.008***
	(28.45)	(34.02)	(21.35)	(23.95)	(28.51)	(24.07)	(27.25)	(22.66)	(22.01)	(21.28)	(18.67)	(13.77)
BSI_Amount_lag5	0.010***	0.013***	0.015***	0.018***	0.015***	0.017***	0.014***	0.014***	0.013***	0.012***	0.009***	0.011***
	(16.61)	(21.52)	(25.64)	(29.99)	(25.08)	(29.04)	(23.81)	(22.93)	(22.48)	(19.34)	(15.09)	(18.62)
EP_Ratio_lag	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.002	-0.002	-0.002	-0.001	-0.002	-0.002
	(-0.35)	(-0.39)	(-0.35)	(-0.30)	(-0.31)	(-0.38)	(-0.25)	(-0.28)	(-0.24)	(-0.15)	(-0.21)	(-0.23)
Turnover_lag	44.889***	-3.709***	5.565***	4.733***	1.758**	0.516	2.671***	1.976***	2.042***	2.180***	2.562***	2.875***
	(62.53)	(-5.14)	(7.67)	(6.51)	(2.41)	(0.71)	(3.65)	(2.70)	(2.78)	(2.96)	(3.48)	(3.90)
Log_Marketvalue_lag	2.088***	0.813***	0.593***	0.636***	0.723***	0.794***	0.868***	0.962***	0.995***	1.028***	1.043***	1.051***
	(44.97)	(17.43)	(12.65)	(13.52)	(15.34)	(16.79)	(18.35)	(20.28)	(20.94)	(21.60)	(21.88)	(22.02)
Constant	-39.687***	-18.304***	-15.346***	-16.194***	-17.658***	-18.804***	-20.040***	-21.526***	-22.089***	-22.645***	-22.928***	-23.081***
	(-53.77)	(-24.69)	(-20.58)	(-21.65)	(-23.56)	(-25.03)	(-26.64)	(-28.56)	(-29.26)	(-29.95)	(-30.27)	(-30.42)
Observations	2,931,240	2,928,258	2,925,160	2,922,143	2,919,092	2,916,072	2,913,048	2,910,033	2,906,996	2,904,007	2,901,023	2,898,045
R-squared	0.061	0.055	0.048	0.046	0.045	0.044	0.044	0.044	0.044	0.044	0.043	0.043
Stock FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table IV Baidu Search Volume Index

This table controls for Baidu Search Volume Index (SVI) as an alternative measure of attention, and run horse-race regressions between attention measures and Baidu Search Volume Index. Panel A focuses on logarithm of Baidu SVI only. Panel B focuses on Community Attention Index, Research Report Attention Index, News Attention Index, and Announcement Attention Index when additional controlling for Baidu Search Volume Index. This table adopts Fama-MacBeth regression.

Panel A Baidu SVI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Return											
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10	T+11
Baidu SVI	1.800***	-0.033	-0.108***	-0.077***	-0.046**	-0.051***	-0.046**	-0.033*	-0.039**	-0.034*	-0.010	-0.015
	(46.93)	(-1.40)	(-4.91)	(-3.64)	(-2.22)	(-2.61)	(-2.39)	(-1.73)	(-2.08)	(-1.85)	(-0.56)	(-0.84)
Return_lag1	-0.045***	0.031***	0.006**	0.001	0.003	-0.009***	0.005**	0.005*	-0.001	0.004*	-0.005**	0.002
	(-14.23)	(9.98)	(2.05)	(0.28)	(1.10)	(-3.44)	(2.00)	(1.92)	(-0.29)	(1.65)	(-2.40)	(0.81)
Return_lag2	-0.051***	0.010***	0.004	0.005*	-0.007***	0.007***	0.006***	0.000	0.006***	-0.004*	0.002	0.003
	(-18.13)	(3.39)	(1.62)	(1.90)	(-2.90)	(2.94)	(2.59)	(0.11)	(2.64)	(-1.68)	(0.91)	(1.12)
Return_lag3	-0.047***	0.003	0.005**	-0.008***	0.006**	0.006***	-0.000	0.006**	-0.004*	0.002	0.002	-0.001
	(-17.74)	(1.13)	(2.08)	(-3.14)	(2.42)	(2.70)	(-0.16)	(2.48)	(-1.90)	(0.74)	(1.02)	(-0.48)
Return_lag4	-0.039***	0.005*	-0.007***	0.006**	0.005**	0.000	0.005**	-0.004*	0.001	0.002	-0.001	-0.001
	(-15.27)	(1.94)	(-2.89)	(2.28)	(2.12)	(0.22)	(2.15)	(-1.77)	(0.62)	(0.81)	(-0.41)	(-0.67)
Return_lag5	-0.046***	-0.008***	0.006**	0.005**	-0.000	0.005**	-0.004*	0.001	0.002	-0.001	-0.001	-0.002
	(-18.64)	(-3.13)	(2.58)	(2.19)	(-0.14)	(2.36)	(-1.74)	(0.64)	(1.01)	(-0.39)	(-0.52)	(-0.77)
EP_Ratio_lag	0.503***	0.004	0.036	0.072	0.047	0.047	0.064	0.027	0.027	0.027	0.039	0.044
	(4.96)	(0.04)	(0.35)	(0.68)	(0.45)	(0.44)	(0.62)	(0.26)	(0.26)	(0.26)	(0.37)	(0.42)
Turnover_lag	13.342***	-5.398***	-3.699***	-2.624***	-2.696***	-2.549***	-1.992***	-1.962***	-1.966***	-1.889***	-1.714***	-1.543***
	(27.88)	(-11.58)	(-8.45)	(-6.15)	(-6.39)	(-6.13)	(-4.82)	(-4.84)	(-4.91)	(-4.72)	(-4.37)	(-3.99)
Log_Marketvalue_lag	0.078***	-0.034***	-0.032***	-0.029***	-0.030***	-0.028***	-0.026***	-0.026***	-0.025***	-0.025***	-0.024***	-0.023***
	(10.61)	(-4.53)	(-4.18)	(-3.81)	(-3.92)	(-3.68)	(-3.38)	(-3.40)	(-3.27)	(-3.25)	(-3.08)	(-3.05)
Constant	-1.710***	0.602***	0.570***	0.526***	0.553***	0.526***	0.473***	0.484***	0.473***	0.470***	0.440***	0.440***
	(-13.10)	(4.48)	(4.23)	(3.86)	(4.06)	(3.85)	(3.49)	(3.55)	(3.45)	(3.43)	(3.21)	(3.20)
Observations	1,801,280	1,801,588	1,801,529	1,801,470	1,801,411	1,801,351	1,801,291	1,801,231	1,801,171	1,801,113	1,801,055	1,800,997
R-squared	0.152	0.101	0.094	0.089	0.085	0.083	0.081	0.079	0.078	0.077	0.076	0.075
Number of groups	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337

Panel B Baidu SVI, and Community Attention Index, News Attention Index, Announcement Attention Index, Research Report Attention Index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Return											
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10	T+11
Community Attention Index	-0.038***	0.035***	-0.048***	-0.046***	-0.021***	-0.022***	-0.019***	-0.012**	-0.010**	-0.016***	-0.016***	-0.007
	(-4.63)	(2.87)	(-7.84)	(-8.22)	(-3.75)	(-4.12)	(-3.68)	(-2.35)	(-2.01)	(-3.30)	(-3.21)	(-1.46)
Research Report Attention Index	0.112***	0.212***	0.043***	0.012***	0.008***	0.007***	0.001	0.002	0.004**	0.005**	0.004**	0.003
	(47.39)	(65.30)	(20.06)	(6.11)	(3.99)	(3.47)	(0.28)	(1.33)	(2.07)	(2.39)	(1.99)	(1.47)
News Attention Index	0.106***	0.190***	0.026***	-0.006***	-0.007***	-0.005***	-0.007***	-0.001	-0.000	-0.002	0.001	-0.004**
	(45.54)	(62.91)	(12.13)	(-3.02)	(-3.78)	(-2.69)	(-3.61)	(-0.62)	(-0.04)	(-1.03)	(0.32)	(-2.42)
Announcement Attention Index	0.008***	0.007***	0.004**	0.001	0.003**	0.004***	0.003*	0.000	0.002	0.002	0.001	0.002
	(4.79)	(4.34)	(2.31)	(0.98)	(2.08)	(2.59)	(1.94)	(0.10)	(1.02)	(1.59)	(0.96)	(1.56)
Baidu SVI	1.504***	-0.646***	-0.154***	-0.029	-0.021	-0.031*	-0.016	-0.022	-0.036**	-0.021	0.002	-0.004
	(43.16)	(-25.21)	(-7.56)	(-1.49)	(-1.11)	(-1.69)	(-0.90)	(-1.22)	(-2.01)	(-1.17)	(0.12)	(-0.23)
Return_lag1	-0.049***	0.014***	0.003	0.001	0.003	-0.009***	0.005**	0.005*	-0.001	0.004	-0.006**	0.002
	(-15.80)	(4.45)	(1.20)	(0.23)	(1.07)	(-3.48)	(2.16)	(1.91)	(-0.39)	(1.64)	(-2.47)	(0.88)
Return_lag2	-0.054***	0.003	0.003	0.005*	-0.007***	0.007***	0.006***	0.000	0.006**	-0.004*	0.002	0.003
	(-19.38)	(1.23)	(1.19)	(1.80)	(-2.91)	(2.93)	(2.60)	(0.01)	(2.58)	(-1.72)	(0.90)	(1.15)
Return_lag3	-0.049***	-0.003	0.004*	-0.008***	0.006**	0.007***	-0.000	0.006**	-0.004*	0.002	0.002	-0.001
	(-18.70)	(-1.14)	(1.71)	(-3.15)	(2.47)	(2.74)	(-0.08)	(2.45)	(-1.92)	(0.76)	(1.06)	(-0.41)
Return_lag4	-0.040***	-0.000	-0.008***	0.006**	0.005**	0.001	0.005**	-0.004*	0.002	0.002	-0.001	-0.001
	(-15.94)	(-0.14)	(-3.21)	(2.38)	(2.19)	(0.25)	(2.21)	(-1.78)	(0.71)	(0.79)	(-0.36)	(-0.60)
Return_lag5	-0.046***	-0.011***	0.006**	0.005**	-0.000	0.005**	-0.004*	0.002	0.002	-0.001	-0.001	-0.001
	(-19.16)	(-4.82)	(2.44)	(2.25)	(-0.13)	(2.41)	(-1.69)	(0.74)	(1.01)	(-0.37)	(-0.45)	(-0.69)
EP_Ratio_lag	0.488***	-0.018	0.040	0.080	0.046	0.047	0.066	0.029	0.035	0.038	0.041	0.052
	(4.88)	(-0.18)	(0.39)	(0.77)	(0.44)	(0.45)	(0.64)	(0.28)	(0.34)	(0.36)	(0.40)	(0.49)
Turnover_lag	12.849***	-6.581***	-3.691***	-2.496***	-2.654***	-2.531***	-1.944***	-1.921***	-1.922***	-1.838***	-1.673***	-1.558***
	(26.72)	(-14.29)	(-8.44)	(-5.88)	(-6.31)	(-6.12)	(-4.73)	(-4.74)	(-4.81)	(-4.61)	(-4.28)	(-4.03)
Log_Marketvalue_lag	0.061***	-0.071***	-0.038***	-0.030***	-0.031***	-0.029***	-0.025***	-0.026***	-0.026***	-0.025***	-0.024***	-0.024***
	(8.35)	(-9.53)	(-5.03)	(-4.02)	(-4.05)	(-3.78)	(-3.29)	(-3.40)	(-3.33)	(-3.26)	(-3.17)	(-3.10)
Constant	-1.404***	1.273***	0.671***	0.544***	0.563***	0.534***	0.459***	0.483***	0.478***	0.468***	0.450***	0.445***
	(-10.94)	(9.43)	(5.00)	(4.01)	(4.15)	(3.92)	(3.38)	(3.52)	(3.48)	(3.42)	(3.27)	(3.24)
Observations	1,801,280	1,801,588	1,801,529	1,801,470	1,801,411	1,801,351	1,801,291	1,801,231	1,801,171	1,801,113	1,801,055	1,800,997
R-squared	0.166	0.142	0.101	0.094	0.091	0.088	0.086	0.084	0.082	0.081	0.081	0.080
Number of groups	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337

Table V Attention and Sentiment

Each stock forum post's sentiment is classified into positive, negative, or neutral, which is obtained from Chinese Research Data Services (CNRDS). The daily totals for each category within a listed company's forum are aggregated. Therefore, the following ratio describes the sentiment of each stock.

$$Pos_Neg_{it} = \frac{Pos_{it} - Neg_{it}}{Pos_{it} + Neg_{it}}$$

Pos_{it} is the number of positive comments for stock i on day t. Neg_{it} is the number of negative comments for stock i on day t.

Panel A focuses on sentiment effect alone. Panel B studies the interaction effect between high sentiment dummy and Community Attention Index, News Attention Index, Announcement Attention Index and Research Report Attention Index respectively. A High_Sentiment_Dummy is equal to 1 if Pos_Neg_{it} is larger than the median daily cut-off value. This table adopts Fama-Macbeth regression. Panel B regression is as follows

$$Ret_{i,t+k} = a0 + \beta 1_{i,t} Attention_{i,t} + \beta 2_{i,t} Attention_{i,t} * High\ Sentiment_{i,t} + \beta 3_{i,t} High\ Sentiment_{i,t} + \beta 4_{i,t} \log_Size_{i,t} + \beta 5_{i,t} EP_ratio_{i,t} + \beta 6_{i,t} Turnover_{i,t} + \beta 7_{i,t} Ret_lag1_{i,t-1} + \beta 8_{i,t} Ret_lag2_{i,t-2} + \beta 9_{i,t} Ret_lag3_{i,t-3} + \beta 10_{i,t} Ret_lag4_{i,t-4} + \beta 11_{i,t} Ret_lag5_{i,t-5} + \varepsilon_{i,t+k}$$

Panel A Sentiment Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Return					
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10
Sentiment	0.776***	0.039***	0.002	-0.001	0.001	-0.007	0.008*	0.001	0.002	0.005	-0.006
	(82.87)	(7.53)	(0.50)	(-0.27)	(0.16)	(-1.46)	(1.65)	(0.24)	(0.50)	(1.09)	(-1.22)
Return_lag1	0.017***	0.001	-0.001	0.000	-0.010***	0.005*	0.003	-0.002	0.002	-0.005**	0.002
	(5.93)	(0.52)	(-0.35)	(0.20)	(-4.30)	(1.91)	(1.33)	(-1.03)	(0.85)	(-2.25)	(0.87)
Return_lag2	-0.000	0.002	0.002	-0.008***	0.007***	0.005**	-0.002	0.004**	-0.003	0.002	0.002
	(-0.07)	(0.85)	(0.97)	(-3.72)	(2.82)	(1.98)	(-0.70)	(1.96)	(-1.59)	(0.96)	(0.86)
Return_lag3	-0.003	0.003	-0.009***	0.006**	0.005**	-0.002	0.004*	-0.003	0.001	0.002	-0.003
	(-1.12)	(1.31)	(-4.03)	(2.47)	(2.15)	(-0.78)	(1.78)	(-1.57)	(0.58)	(0.90)	(-1.25)
Return_lag4	-0.000	-0.008***	0.005**	0.003	-0.001	0.004*	-0.003	0.001	0.001	-0.003	-0.002
	(-0.09)	(-3.79)	(2.22)	(1.40)	(-0.52)	(1.80)	(-1.58)	(0.64)	(0.63)	(-1.28)	(-0.88)
Return_lag5	-0.010***	0.006***	0.004	-0.001	0.004*	-0.003	0.001	0.002	-0.003	-0.002	-0.002
	(-4.76)	(2.72)	(1.61)	(-0.55)	(1.94)	(-1.32)	(0.63)	(0.85)	(-1.45)	(-0.80)	(-1.15)
EP_Ratio_lag	-0.113	-0.059	-0.019	-0.094	-0.058	-0.049	-0.070	-0.068	-0.075	-0.039	-0.037
	(-1.01)	(-0.52)	(-0.17)	(-0.82)	(-0.51)	(-0.43)	(-0.62)	(-0.60)	(-0.67)	(-0.34)	(-0.32)

Turnover_lag	-4.398***	-4.103***	-2.853***	-2.767***	-2.724***	-2.203***	-2.053***	-2.049***	-1.940***	-1.698***	-1.535***
	(-10.10)	(-9.96)	(-7.19)	(-7.08)	(-7.13)	(-5.77)	(-5.47)	(-5.54)	(-5.28)	(-4.71)	(-4.30)
Log_Marketvalue_lag	-0.047***	-0.033***	-0.029***	-0.028***	-0.027***	-0.025***	-0.025***	-0.025***	-0.024***	-0.023***	-0.022***
	(-6.31)	(-4.38)	(-3.76)	(-3.68)	(-3.53)	(-3.22)	(-3.23)	(-3.19)	(-3.05)	(-2.95)	(-2.88)
Constant	0.653***	0.584***	0.513***	0.519***	0.500***	0.451***	0.459***	0.462***	0.443***	0.423***	0.422***
	(4.96)	(4.34)	(3.78)	(3.81)	(3.67)	(3.33)	(3.36)	(3.36)	(3.23)	(3.08)	(3.07)
Observations	2,631,137	2,631,050	2,630,963	2,630,875	2,630,788	2,630,699	2,630,611	2,630,522	2,630,436	2,630,350	2,630,265
R-squared	0.107	0.079	0.076	0.073	0.071	0.069	0.068	0.067	0.066	0.065	0.065
Number of groups	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337

Panel B Interaction Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Return					
	T	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8	T+9	T+10
Community Attention Index	-0.375***	-0.104***	-0.059***	-0.031***	-0.023***	-0.023***	-0.018***	-0.017***	-0.015***	-0.016***	-0.008
	(-29.02)	(-15.02)	(-8.98)	(-5.12)	(-3.88)	(-3.89)	(-3.17)	(-3.00)	(-2.68)	(-2.84)	(-1.40)
Research Report Attention Index	0.140***	0.036***	0.010***	0.008***	0.004**	0.000	0.002	0.006***	0.004*	0.003	0.002
	(45.98)	(15.83)	(4.82)	(3.66)	(2.07)	(0.23)	(1.18)	(2.89)	(1.70)	(1.55)	(0.85)
News Attention Index	0.115***	0.002	-0.010***	-0.007***	-0.005**	-0.004	0.000	-0.002	-0.002	0.001	-0.003
	(38.93)	(0.96)	(-4.17)	(-2.91)	(-2.11)	(-1.60)	(0.18)	(-1.03)	(-0.77)	(0.29)	(-1.44)
Announcement Attention Index	-0.007***	-0.002	0.001	0.004**	0.004**	0.002	-0.000	0.003*	0.001	0.003	0.002
	(-3.78)	(-0.84)	(0.32)	(2.24)	(2.21)	(1.10)	(-0.01)	(1.74)	(0.79)	(1.43)	(0.99)
Community Attention Index_High Sentiment	0.713***	0.073***	0.007	0.006	-0.004	-0.005	0.006	0.002	-0.004	0.004	0.001
	(64.58)	(9.99)	(1.07)	(0.95)	(-0.56)	(-0.83)	(0.97)	(0.27)	(-0.59)	(0.61)	(0.17)
Research Report Attention Index_High Sentiment	0.140***	0.021***	0.004	-0.001	0.002	0.000	0.000	-0.004	-0.001	-0.001	-0.001
	(42.46)	(8.53)	(1.41)	(-0.22)	(0.80)	(0.14)	(0.05)	(-1.46)	(-0.30)	(-0.51)	(-0.51)
News Attention Index_High Sentiment	0.120***	0.040***	0.006*	-0.001	-0.001	-0.008***	-0.006**	0.003	0.000	-0.004	-0.003
	(33.18)	(12.95)	(1.85)	(-0.38)	(-0.21)	(-2.65)	(-2.03)	(0.90)	(0.05)	(-1.34)	(-1.17)
Announcement Attention Index_High Sentiment	0.017***	0.004*	0.003	0.001	-0.002	0.000	0.001	-0.004*	0.002	-0.001	-0.001
	(6.53)	(1.66)	(1.15)	(0.33)	(-0.72)	(0.06)	(0.36)	(-1.78)	(0.69)	(-0.21)	(-0.39)
High Sentiment Dummy	0.713***	0.021***	-0.001	0.003	-0.003	-0.003	0.012**	0.006	0.004	0.010**	0.001
	(79.37)	(4.09)	(-0.15)	(0.65)	(-0.61)	(-0.62)	(2.42)	(1.19)	(0.86)	(2.04)	(0.11)
Return_lag1	-0.015***	-0.002	-0.000	0.001	-0.009***	0.006**	0.003	-0.002	0.002	-0.005**	0.002
	(-5.45)	(-0.80)	(-0.08)	(0.31)	(-4.02)	(2.39)	(1.41)	(-1.03)	(0.98)	(-2.20)	(1.09)
Return_lag2	-0.016***	0.000	0.003	-0.008***	0.007***	0.005**	-0.001	0.004**	-0.003	0.002	0.002
	(-6.13)	(0.15)	(1.22)	(-3.57)	(2.95)	(2.21)	(-0.62)	(1.99)	(-1.48)	(1.06)	(1.05)
Return_lag3	-0.015***	0.002	-0.008***	0.006***	0.005**	-0.001	0.004*	-0.003	0.001	0.002	-0.002
	(-6.08)	(0.89)	(-3.77)	(2.67)	(2.38)	(-0.42)	(1.82)	(-1.52)	(0.66)	(1.00)	(-1.05)
Return_lag4	-0.010***	-0.009***	0.006**	0.004	-0.001	0.004**	-0.003	0.002	0.002	-0.002	-0.001
	(-4.56)	(-4.19)	(2.50)	(1.62)	(-0.33)	(2.03)	(-1.48)	(0.74)	(0.71)	(-1.13)	(-0.64)
Return_lag5	-0.018***	0.006***	0.004*	-0.001	0.004**	-0.002	0.002	0.002	-0.003	-0.001	-0.002
	(-8.27)	(2.59)	(1.86)	(-0.39)	(2.07)	(-1.13)	(0.81)	(0.94)	(-1.31)	(-0.60)	(-0.96)

EP_Ratio_lag	-0.154	-0.055	-0.004	-0.086	-0.050	-0.039	-0.073	-0.068	-0.072	-0.041	-0.034
	(-1.42)	(-0.49)	(-0.03)	(-0.76)	(-0.44)	(-0.35)	(-0.65)	(-0.60)	(-0.65)	(-0.37)	(-0.30)
Turnover_lag	-6.321***	-3.928***	-2.433***	-2.567***	-2.574***	-1.997***	-1.915***	-1.952***	-1.809***	-1.590***	-1.470***
	(-14.90)	(-9.59)	(-6.16)	(-6.56)	(-6.75)	(-5.21)	(-5.08)	(-5.25)	(-4.89)	(-4.38)	(-4.07)
Log_Marketvalue_lag	-0.074***	-0.039***	-0.029***	-0.029***	-0.027***	-0.024***	-0.025***	-0.025***	-0.024***	-0.024***	-0.022***
	(-10.09)	(-5.13)	(-3.81)	(-3.74)	(-3.59)	(-3.14)	(-3.25)	(-3.27)	(-3.09)	(-3.04)	(-2.92)
Constant	0.883***	0.659***	0.510***	0.517***	0.505***	0.440***	0.454***	0.466***	0.444***	0.430***	0.423***
	(6.73)	(4.92)	(3.77)	(3.82)	(3.72)	(3.24)	(3.32)	(3.39)	(3.24)	(3.12)	(3.09)
Observations	2,631,137	2,631,050	2,630,963	2,630,875	2,630,788	2,630,699	2,630,611	2,630,522	2,630,436	2,630,350	2,630,265
R-squared	0.163	0.089	0.084	0.080	0.078	0.076	0.075	0.073	0.072	0.072	0.071
Number of groups	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337

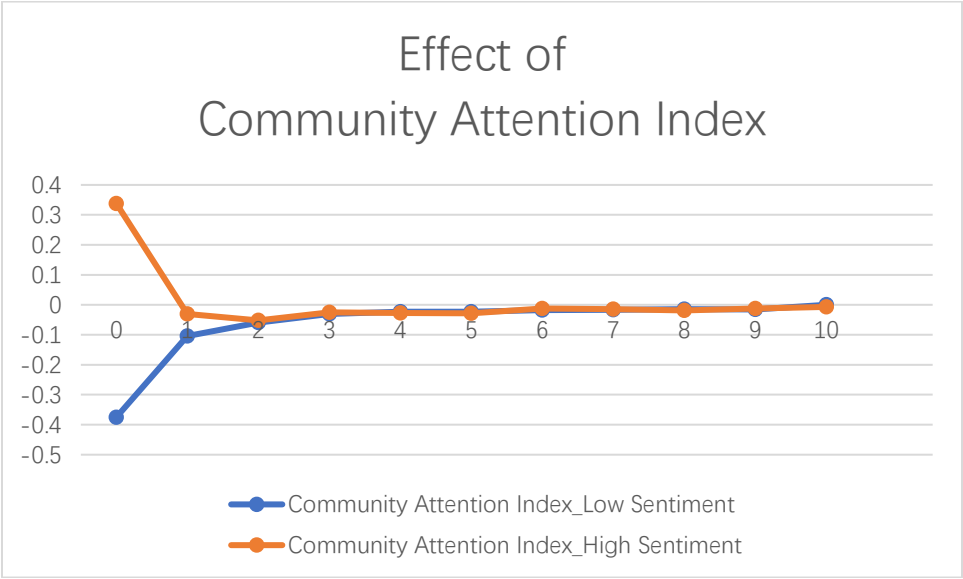


Figure 3a Community Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the Fama-MacBeth regression coefficients of Community Attention Index on stock returns from Table V Panel B, separately for high sentiment days (above median) and low sentiment days (below median). Specifically, the effect for low sentiment periods at horizon k (x-axis) equals $\beta 1_{t+j}$. The effect for high sentiment periods at horizon k (x-axis) equals $\beta 1_{t+j} + \beta 2_{t+j}$, where $\beta 1_{t+j}$ is the coefficient for attention measures on day T+j and $\beta 2_{t+j}$ is the coefficient for attention*high sentiment interaction terms on day T+j from Table V Panel B. The figure illustrates how unsophisticated retail attention exhibits opposite return predictability patterns depending on market sentiment—positive followed by reversal during high sentiment periods versus negative during low sentiment periods.

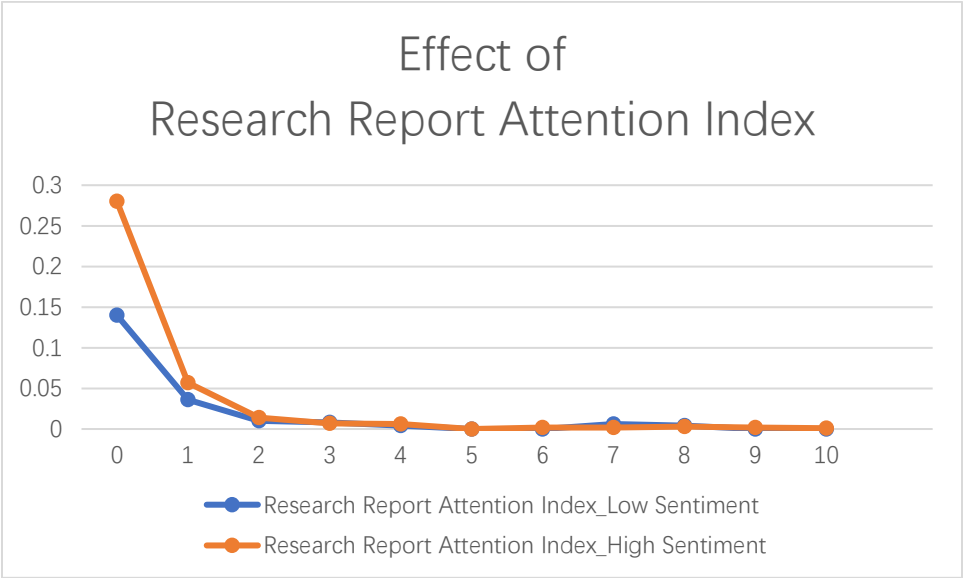


Figure 3b Research Report Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the Fama-MacBeth regression coefficients of Research Report Attention Index on stock returns from Table V Panel B, separately for high sentiment days and low sentiment days. The figure shows that sophisticated attention positively predicts returns under both sentiment regimes, with amplified effects during high sentiment periods and no evidence of return reversal.

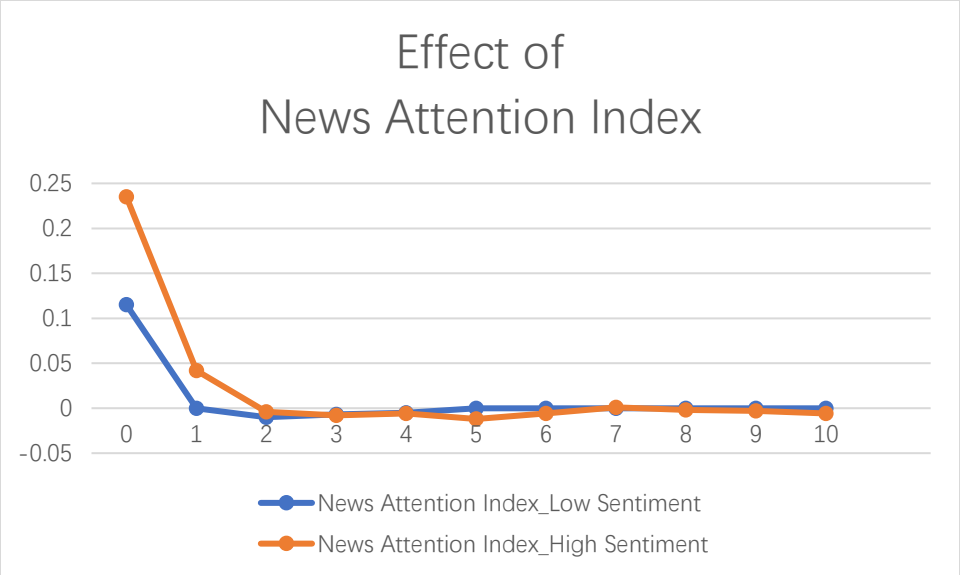


Figure 3c News Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the Fama-MacBeth regression coefficients of News Attention Index on stock returns from Table V Panel B, separately for high sentiment days and low sentiment days. The figure illustrates the initial positive return predictability followed by partial reversal under both sentiment regimes, consistent with short-lived informational cascades.

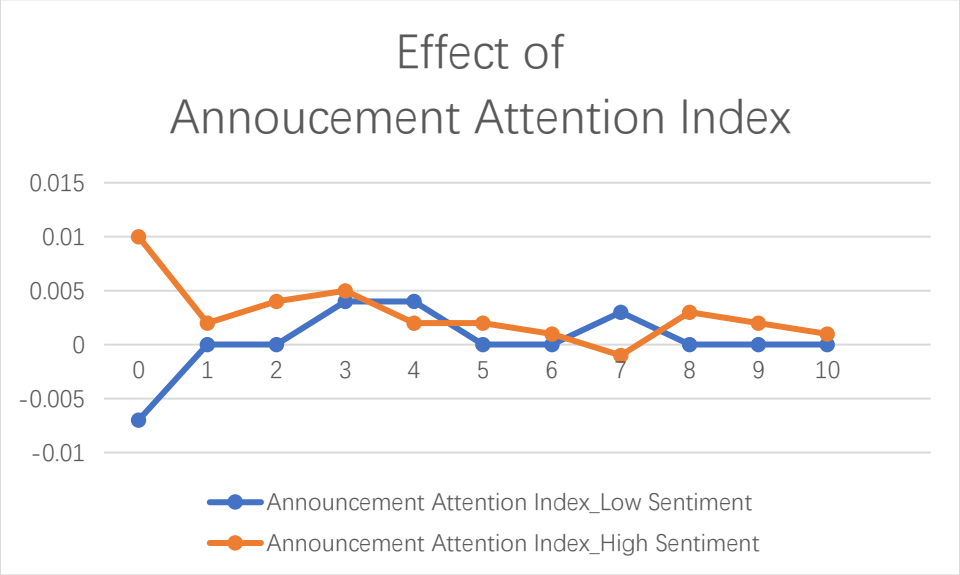


Figure 3d Announcement Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the Fama-MacBeth regression coefficients of Announcement Attention Index on stock returns from Table V Panel B, separately for high sentiment days and low sentiment days. The figure reveals delayed positive effects during low sentiment periods (T+3 and T+4) consistent with post-earnings announcement drift, versus immediate effects during high sentiment periods.

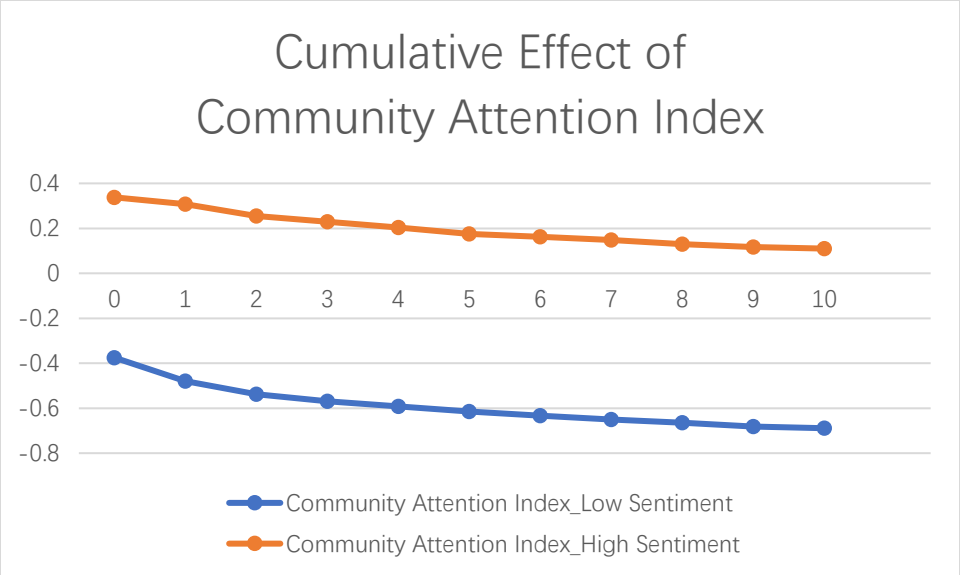


Figure 4a Cumulative Community Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the cumulative sum of Fama-MacBeth regression coefficients of Community Attention Index from Table V Panel B, separately for high sentiment and low sentiment days. Specifically, the cumulative effect for low sentiment periods at horizon k (x-axis) equals $\sum_{j=0}^{k} \beta 1_{t+j}$, where $\beta 1_{t+j}$ is the coefficient of attention for day T+j from Table V Panel B. The cumulative effect for high sentiment periods at horizon k (x-axis) equals $\sum_{j=0}^{k} \beta 1_{t+j} + \sum_{j=0}^{k} \beta 2_{t+j}$, where $\beta 1_{t+j}$ is the coefficient for attention measures on day T+j and where $\beta 2_{t+j}$ is the coefficient for attention*high sentiment interaction terms on day T+j from Table V Panel B. The figure reveals a striking contrast: positive cumulative effects during high sentiment periods versus negative cumulative effects during low sentiment periods, demonstrating that unsophisticated investor attention is highly influenced by market sentiment.

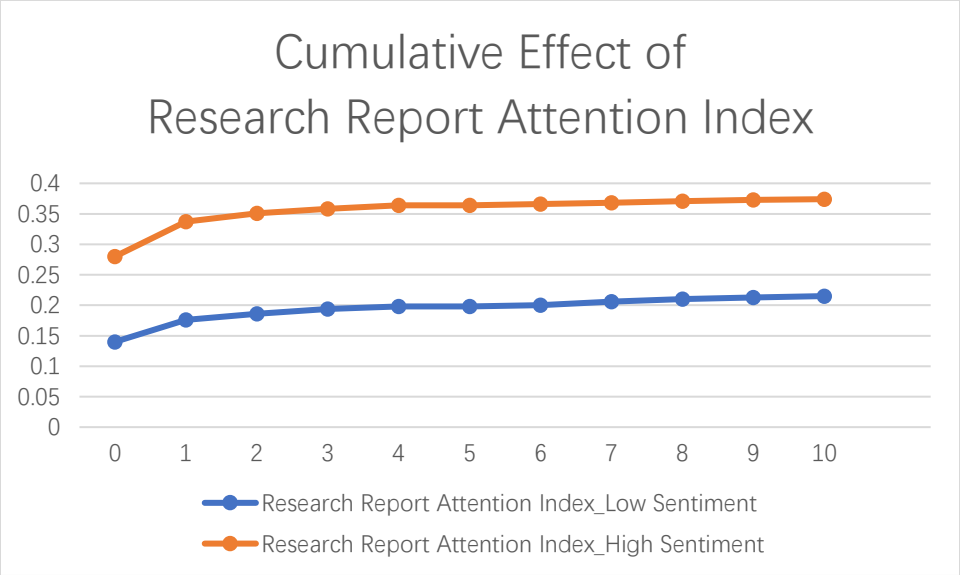


Figure 4b Cumulative Research Report Attention Index Coefficients Under High vs. Low Sentiment Periods

This figure plots the cumulative sum of Fama-MacBeth regression coefficients of Research Report Attention Index from Table V Panel B, separately for high sentiment and low sentiment days. The figure shows sustained positive cumulative effects under both sentiment regimes, with higher cumulative returns during high sentiment periods, confirming that sophisticated attention captures fundamental information regardless of market mood.

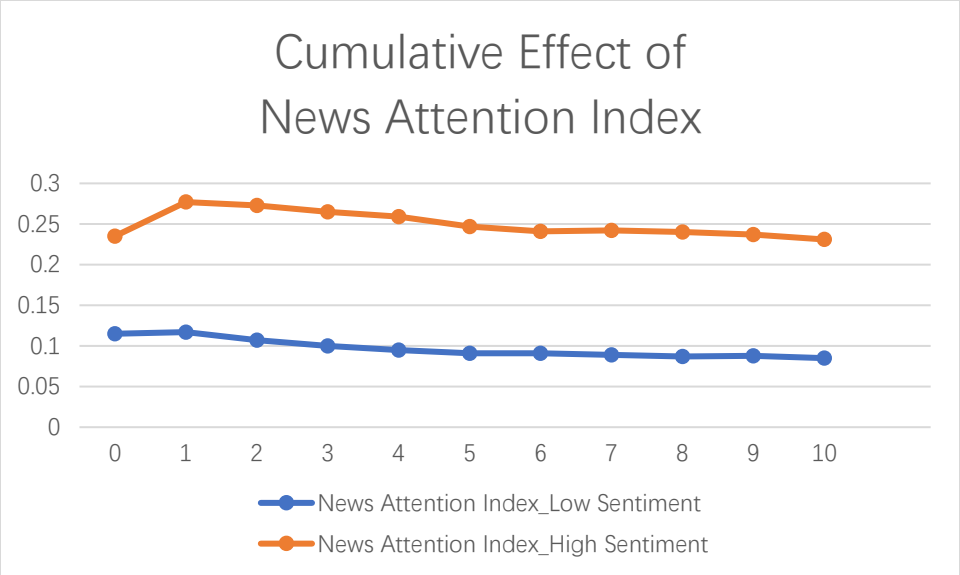


Figure 4c Cumulative News Attention Index Coefficients Under High vs. Low Sentiment Periods
This figure plots the cumulative sum of Fama-MacBeth regression coefficients of News Attention Index from Table V Panel B, separately for high sentiment and low sentiment days. The figure illustrates positive cumulative effects that plateau and partially reverse, indicating that news attention contains both informational and noise components.

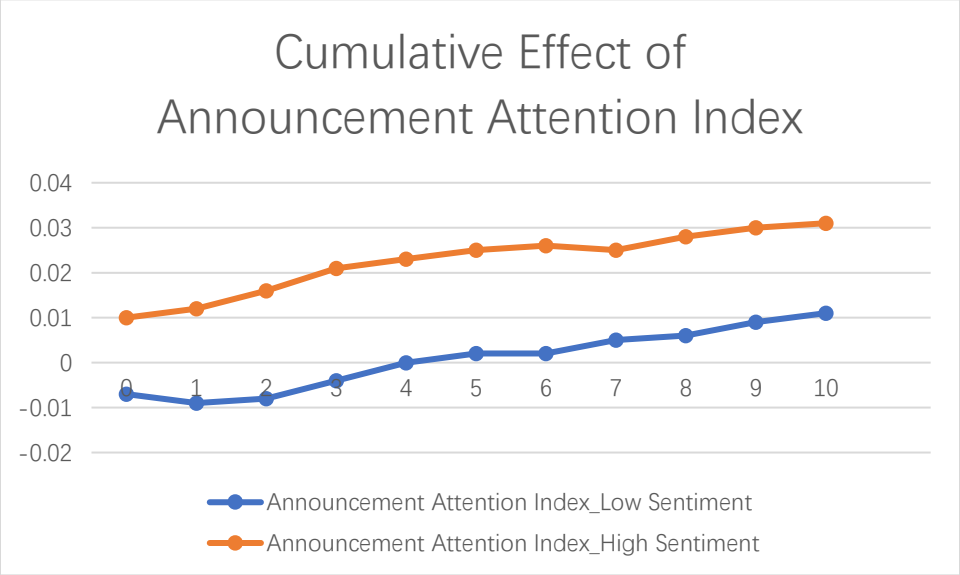


Figure 4d Cumulative Announcement Attention Index Coefficients Under High vs. Low Sentiment Periods
This figure plots the cumulative sum of Fama-MacBeth regression coefficients of Announcement Attention Index from Table V Panel B, separately for high sentiment and low sentiment days. The figure shows moderate positive cumulative effects, with delayed price discovery evident during low sentiment periods, consistent with initial underreaction to corporate announcements followed by gradual correction.

Appendix Variable Definition

Variable Name	Description	Measurement	Source
Community Attention Index	Unsophisticated attention proxy	Abnormal attention is calculated following Da, Engelberg, and Gao (2011), using a 30-day median window to capture normal attention levels. Log(Abnormal attention) = log(Attention_t) - log(median of prior 30 days); Community Attention Index is measured by number of posts and replies on Guba EastMoney stock message board.	EastMoney app
Research Report Attention Index	Sophisticated attention proxy	Abnormal attention is calculated following Da, Engelberg, and Gao (2011), using a 30-day median window to capture normal attention levels. Log(Abnormal attention) = log(Attention_t) - log(median of prior 30 days); Research Report Attention Index is measured by number of reads and comments on professional research reports.	EastMoney app
News Attention Index	Attention to news	Abnormal attention is calculated following Da, Engelberg, and Gao (2011), using a 30-day median window to capture normal attention levels. Log(Abnormal attention) = log(Attention_t) - log(median of prior 30 days); News Attention Index is measured by number of news article reads and comments for a given stock.	EastMoney app
Announcement Attention Index	Attention to corporate announcements	Abnormal attention is calculated following Da, Engelberg, and Gao (2011), using a 30-day median window to capture normal attention levels. Log(Abnormal attention) = log(Attention_t) - log(median of prior 30 days); Announcement Attention Index is measured by number of reads and comments on corporate announcements.	EastMoney app
Stock Return	Daily stock return	Percentage change in stock price	CSMAR
BSI_Amount_Unsophisticated	Buy-sell imbalance for unsophisticated retail investors	(Buy_Amount - Sell_Amount) / (Buy_Amount + Sell_Amount); small and middle trades (<100,000 shares or <CNY 200,000)	EastMoney moneyflow
BSI_Amount_Sophisticated	Buy-sell imbalance for sophisticated retail and institutional investors	(Buy_Amount - Sell_Amount) / (Buy_Amount + Sell_Amount); large and extra-large trades (≥100,000 shares or ≥CNY 200,000)	EastMoney moneyflow
Baidu SVI	Alternative attention proxy	Log of abnormal Search Volume Index from Baidu search engine.	Baidu
Sentiment (Pos_Neg)	Investor sentiment from stock forum posts	(Pos_it - Neg_it) / (Pos_it + Neg_it); where Pos = number of positive comments, Neg = number of negative comments	CNRDS
High Sentiment Dummy	Indicator for high sentiment periods	=1 if Pos_Neg[t-1] is above median daily cutoff	CNRDS
EP_Ratio	Earnings-to-price ratio	Earnings per share / price per share	CSMAR
Turnover	Stock liquidity	Trading volume / shares outstanding	CSMAR
Log_Marketvalue	Firm size	Natural logarithm of market value (in thousands of CNY)	CSMAR
Return_lag1 to lag5	Return momentum controls	Five lags of daily stock return	CSMAR