Sticky Attention and Stock Returns*

Liyun Chen
UNSW Business School

Diego García
University of Colorado at Boulder

Alexandre Jeanneret
UNSW Business School

Benjamin Loos
UNSW Business School

August 26, 2026

* We are very grateful for comments and suggestions from Zhi Da, Cesare Fracassi, Amit Goyal, Jun Liu (discussant), Jason Tian (discussant), Albert Wang (discussant), Xiaoyan Zhang, and participants of 2025 Sydney Banking and Financial Stability Conference, 2025 FIRN UQ Asset Management Meeting, 2026 CityUHK International Finance Conference, 2026 Waseda Summer Workshop in Finance, and seminar participants at HEC Lausanne, Penn State University, University of Colorado–Boulder, and University of North Carolina at Chapel Hill. Chen is from UNSW Business School; e-mail: liyun.chen@unsw.edu.au; website: https://www.unsw.edu.au/hdr/liyun-chen. García is from University of Colorado at Boulder; e-mail: diego.garcia@colorado.edu; website: https://leeds-faculty.colorado.edu/garcia. Jeanneret (corresponding author) is from UNSW Business School; e-mail: a.jeanneret@unsw.edu.au; website: www.alexandrejeanneret.net. Loos is from UNSW Business School; e-mail: b.loos@unsw.edu.au; website: www.unsw.edu.au/staff/benjamin-loos. All errors are our own.

Sticky Attention and Stock Returns

ABSTRACT

When a stock retains an unusually large share of the previous day's viewers, an attention shock is associated with a same-day return of over 112 basis points that fully reverses within forty trading days; when retention is unusually low, attention barely moves prices. We document this contrast using individual browsing records from China's largest investment platform, which identify the same viewer across days and allow us to measure attention stickiness: the fraction of the previous day's viewers who return today, a dimension aggregate attention proxies do not capture. The level of stickiness predicts lower future returns, cumulating to 160 basis points over forty days. In the cross-section, stocks both rarely viewed and rarely revisited earn 2.2% per month more than heavily monitored, high-stickiness stocks. The composition of retail attention, not merely its level, is a first-order predictor of short-horizon returns; the evidence is more consistent with behavioral reinforcement than with rational learning.

JEL Classification Numbers: G12, G14, G41

Keywords: Investor attention, Sticky attention, Return predictability, Price pressure.

1 Introduction

Investor attention is usually treated as a quantity: the number of searches, views, or queries directed at a stock. This treatment overlooks a basic distinction. The same measured attention may reflect a stable audience monitoring the stock day after day, or a succession of investors arriving for the first time. The two forms of attention can generate very different demand: returning investors may trade after repeated exposure to the same stock, whereas a rotating viewer base may reflect more heterogeneous information and trading intentions. Whether attention moves prices may therefore depend on the composition and persistence of a stock's audience, not merely on how much attention it receives.

Aggregate attention measures cannot make this distinction because they collapse the viewer set to a count. Google search volume (Da et al., 2011), Bloomberg terminal queries (Ben-Rephael et al., 2017), watchlist additions (Levy et al., 2024), news readership (Cookson et al., 2024), and platform rankings such as Robinhood's "Top Mover" list (Barber et al., 2022) all assign the same value to two episodes with identical aggregate attention. One episode may be generated by a stable viewer base, the other by a largely different set of investors. A stock viewed by one thousand investors on two consecutive days may be monitored by the same thousand or by two thousand different ones; aggregate attention proxies record the two cases identically. As a result, aggregate attention measures cannot test whether the price response to attention depends on the persistence of the audience generating it.

We use individual browsing records from a large Chinese investment platform that identify which users view which stocks on each day.[1] These data allow us to construct a model-free measure of viewer-base persistence, which we call *attention stickiness*. The measure $STK_{i,t}$ is the fraction of stock i's previous-day viewers who return the following day. It is distinct from persistence in aggregate view counts: a stock can receive a stable number of views while its viewer base turns over almost completely.

The central result is that attention has no single price effect: its effect depends on who is paying attention. Attention shocks in our sample are followed by the price-pressure-and-reversal pattern documented for U.S. markets, but the average masks two very different episodes. The estimated same-day price response to abnormal attention exceeds 112 basis points when viewer retention is above the median, but is close to zero when retention is below the median. The high-retention response is followed

[1] The platform, *RoyalFlush (Tong-Hua-Shun)*, serves more than 19 million active weekly users; our sample covers 3,144 A-share Main Board stocks from January 2018 to December 2024. According to the 2023 report by China Securities Depository and Clearing Corporation Limited (CSDC), individual investors account for over 90% of traders in the Chinese stock market and contribute over 60% of trading volume. This retail-dominated structure makes China a useful setting: channels operating through retail demand are more likely to be detectable.

by substantial return reversal over subsequent weeks. Aggregate attention measures therefore combine episodes with very different implications for prices. We refer to these as sticky (high-retention) and transitory (low-retention) attention states.

Viewer-base persistence also predicts returns independently of any concurrent attention shock. Conditional on attention, higher viewer retention is followed by lower subsequent returns, and this direct level effect becomes the dominant within-stock predictor at multi-week horizons. Persistence therefore matters through two related but distinct margins: it changes the immediate price response to an attention shock and predicts how prices evolve afterward. More broadly, attention episodes with the same aggregate intensity can have very different price consequences depending on whether they are generated by a stable audience or a rotating viewer base. Aggregate attention measures conceal this distinction because they record how much attention a stock receives, but not the persistence of the audience generating it.

A complementary pattern emerges in the cross-section, using a different source of variation. Sorting stocks monthly on prior-month average attention and viewer retention produces a 5×5 grid in which the largest return differences occur when both dimensions are low or both are high. Stocks that are rarely viewed and rarely revisited earn approximately 2.2 percentage points per month more than stocks that are heavily viewed and have high viewer retention. The spread remains large after adjustment for standard Chinese equity factors and is consistent with a deeper investor-recognition gap for stocks that attract little attention and fail to develop a stable viewer base (Merton, 1987). This evidence complements, but is distinct from, the within-stock panel results: the panel analysis captures changes in a stock's attention composition, whereas the portfolio analysis captures persistent differences in the average monitoring that stocks receive.

Contributions. This paper makes three contributions. First, it identifies viewer-base persistence as a previously unmeasured dimension of investor attention at the stock-audience level. Aggregate measures record how much attention a stock receives, but not whether the same investors continue to monitor it. Second, it shows that attention has a state-dependent price effect: aggregate attention is associated with substantially different price dynamics depending on viewer retention. This result implies that aggregate attention measures can conceal economically important heterogeneity across attention episodes. Third, the evidence connects this new dimension of attention to limits to arbitrage and investor recognition. The effects are strongest where retail investors are more likely to be marginal traders and arbitrage is more constrained, while persistent differences in attention composition are associated

with longer-horizon cross-sectional return differences.

Interpretation. Two interpretations generate different predictions for how viewer-base persistence affects prices. Under behavioral demand concentration, repeated monitoring may produce more coordinated temporary demand, leading to a stronger contemporaneous price response and subsequent reversal. Under rational information aggregation, returning viewers may possess richer information and improve price accuracy, implying that the price response should be more persistent and exhibit less reversal. The two forces may operate simultaneously; the empirical question is which dominates in the observed price response. Because the data observe browsing rather than trading, we interpret the results as reduced-form evidence consistent with behavioral demand concentration, not as direct identification of a particular mechanism. Section 2 develops these predictions.

Identification. Our strategy relies on reduced-form panel regressions with firm and time fixed effects, not on exogenous variation in attention composition. We cannot rule out that persistent time-varying firm characteristics jointly influence STK and future returns. Three features mitigate this concern. The included observables explain only a limited share of the day-to-day variation in STK. The sign structure (a positive contemporaneous association followed by negative forward returns) is more consistent with temporary demand than with a persistent confound. The estimates remain similar after removing the component of STK linearly related to aggregate attention, firm characteristics, and news (Section 4.3); this exercise does not establish exogeneity, but it reduces concerns that the baseline results simply reflect observable differences in attention, liquidity, or information arrival. The pattern also appears after aggregating attention and returns to equally weighted industries, suggesting that it is not entirely driven by a small number of individual stocks.

Related literature. Our work builds on three strands of the attention literature. The first studies *aggregate attention and asset prices*, from indirect proxies such as trading volume and extreme returns (Gervais et al., 2001; Barber and Odean, 2008) to the direct measures discussed above.[2] This literature documents that attention episodes are followed by short-term price pressure and reversal; our contribution is to show that the strength of this pattern depends on a dimension of attention these proxies cannot record. Closest to our emphasis on temporal structure, Da et al. (2014) show that the continuity

[2]Other indirect proxies include news and media coverage (Tetlock, 2011; Yuan, 2015), market index performance (Hou, 2009; Li and Yu, 2012; Vozlyublennaia, 2014), and firms' advertising expenditure (Chemmanur and Yan, 2009). Search behavior also reveals sentiment: Da et al. (2015) aggregate household search terms into a daily index that predicts short-term return reversals.

of the *information* flow shapes underreaction through limited attention; we measure the continuity of the audience itself rather than of the signal.

The second strand studies *heterogeneity within attention*. Barber and Odean (2008) show that attention-constrained retail investors are net buyers of salient stocks. Levy et al. (2024) show that investors' information sets determine whether attention increases or decreases their returns. Da et al. (2025) find that institutional attention predicts positive returns while retail attention predicts negative returns, and Ben-Rephael et al. (2021) show that anticipated information consumption by institutions is associated with a return premium.[3] Our decomposition adds a within-retail dimension: part of the negative retail-attention predictability documented by Da et al. (2025) may reflect differences in the persistence of the retail viewer base. Recent work linking attention to CAPM beta shifts (Andrei et al., 2023) and attention spillovers (Chen et al., 2023) uses aggregate measures and therefore cannot distinguish sticky from transitory attention.

The third strand observes *attention at the individual level*. Sicherman et al. (2016) track logins to retirement accounts, Gargano and Rossi (2018) track web activity within a brokerage platform and find that more attentive investors earn higher returns, and Levy et al. (2024) observe watchlist additions. These studies observe individual investors' attention or activity. Our data instead allow us to characterize the audience for a stock as a whole and to measure how much that audience overlaps with the previous day's audience. The contribution is therefore not simply to observe attention at the individual level, but to show that the persistence of a stock's viewer base is itself an economically relevant dimension of attention.

The remainder of the paper proceeds as follows. Section 2 develops the conceptual framework and states four empirical predictions. Section 3 describes the data and sample. Section 4 presents the main effects of attention stickiness. Section 5 presents portfolio evidence. Section 6 presents cross-sectional tests, industry-level analysis, and robustness checks. Section 7 concludes.

2 Conceptual Framework and Empirical Predictions

Whether viewer-base persistence matters for prices depends on how a persistent audience differs from a rotating one. STK records overlap in a stock's viewers across days; it does not observe trading, order flow, beliefs, or the direction of investors' positions. Holding aggregate attention fixed, high

[3] Ben-Rephael et al. (2017) develop a measure of abnormal institutional attention using Bloomberg terminal queries; Ben-Rephael et al. (2022) show that institutional attention leads retail attention around SEC disclosures; and Kempf et al. (2017) document real effects of institutional distraction.

STK identifies a relatively stable audience, whereas low STK identifies an audience with relatively little overlap with the previous day's viewers. A stable audience may generate more correlated trading intentions, while a rotating audience may be associated with more heterogeneous trading responses. We consider two interpretations of this distinction, behavioral demand concentration and rational information aggregation, and derive four reduced-form predictions.

2.1 Behavioral Demand Concentration

Under the behavioral-demand-concentration interpretation, which we also refer to as behavioral reinforcement, repeated exposure to the same stock may strengthen existing beliefs, increase familiarity, or make investors' trading intentions more correlated.[4] The key implication is not that every returning viewer buys, but that a persistent audience may generate more coordinated trading intentions than a rotating audience. Because attention-constrained retail investors are disproportionately likely to buy rather than sell attention-grabbing stocks (Barber and Odean, 2008), greater coordination among returning viewers can translate into stronger positive net demand. Conditional on aggregate attention, high viewer retention should be associated with a larger contemporaneous price response. If the resulting demand is temporary, the price response should subsequently reverse as the attention episode fades. Viewer retention may also capture coordinated demand when aggregate attention is close to its normal level: if a stable audience sustains temporary price pressure independently of an abnormal attention shock, higher STK should directly predict lower subsequent returns. The individual psychological forces that could produce this pattern are not separately identifiable in our data; we therefore treat them as alternative sources of a common behavioral interpretation.

2.2 Rational Information Aggregation

Under a rational-information-aggregation interpretation, which we refer to as rational learning for brevity, returning viewers may possess a richer history of signals and process information more accurately (Andrei and Hasler, 2015; Peng, 2005). Viewer persistence should then improve the informational quality of demand. High-retention episodes may still generate a contemporaneous price response, but that response should contain more information and therefore exhibit less subsequent reversal. Rational learning does not necessarily imply positive subsequent returns or a smaller same-day price movement; its sharper

[4] Several psychological forces could operate jointly: commitment escalation, in which each revisit deepens an implicit stake and raises the propensity to act on it (Staw, 1976; Thaler, 1980); confirmation bias, in which repeated exposure amplifies the tendency to interpret ambiguous signals as supportive of a prior view (Rabin and Schrag, 1999); and familiarity, which compresses perceived risk (Huberman, 2001).

implication is that persistence should not generate the negative post-event return pattern predicted by temporary behavioral demand.

The two forces may operate simultaneously; the empirical question is which dominates in the observed price response. Figure 1 summarizes the two interpretations.

FIGURE 1 ABOUT HERE

2.3 Empirical Predictions

The two interpretations generate different predictions for the relative magnitude and persistence of the price response. The empirical analysis focuses on whether high viewer retention amplifies contemporaneous price pressure and subsequent reversal, rather than on whether it produces positive or negative returns at every horizon. We state four predictions that organize the analysis.

P1. (Contemporaneous price response) Conditional on abnormal attention, the estimated same-day return response is larger when viewer retention is high than when it is low. This prediction follows from the view that a persistent audience generates more coordinated demand. Under rational information aggregation, returning viewers may also move prices on the event day, but through information rather than temporary demand; the interpretations are distinguished primarily by whether the response persists (P2).

P2. (Transitory high-retention response) If the stronger contemporaneous response in the high-retention state reflects temporary demand, the cumulative price response associated with a high-retention attention episode should decline over subsequent trading days. Under rational information aggregation, a price response generated by better-informed returning viewers should be more persistent and exhibit less reversal.

P3. (Independent predictive content) Conditional on aggregate attention, higher viewer retention predicts lower subsequent returns even after accounting for its role in moderating the response to the attention shock. This prediction concerns the direct predictive content of viewer composition. Under rational information aggregation, viewer retention should not generate a systematic negative post-event return.

P4. (Retail clientele and limits to arbitrage) The high- versus low-retention reversal gap should be larger where retail investors are more likely to be marginal traders and where arbitrage is more constrained, such as among low-price stocks and firms with low institutional ownership. Such patterns would be consistent with viewer persistence operating through retail demand and limited arbitrage, although they would not rule out all informational mechanisms.

Empirically, P1 concerns the contemporaneous interaction between abnormal attention and viewer retention. P2 concerns whether the total predicted return associated with a high-retention attention episode declines after its contemporaneous peak. P3 concerns the direct coefficient on viewer retention, conditional on attention and the interaction term. P4 concerns whether the high- versus low-retention reversal gap is larger in segments where retail investors are more likely to be marginal traders and arbitrage is more constrained.

These predictions concern the reduced-form price implications of viewer-base persistence. They do not identify whether repeated viewing changes beliefs, increases familiarity, or generates buying rather than selling. Because the data do not link browsing records to individual trades, the evidence characterizes how price dynamics differ between persistent and rotating audiences, while the behavioral interpretation remains indirect. Sections 4–6 present the evidence.

3 Data and Sample

3.1 RoyalFlush Platform

Our primary data source is *RoyalFlush* (*Tong-Hua-Shun*), China's largest investment information and trading platform, which provides comprehensive data on A-share listed firms.[5] The platform combines fundamental and technical research, real-time market data, and integrated brokerage services, making it the dominant retail-investor tool in China. Its 2024 annual report records more than 19 million active users per week, a base large enough to treat its browsing data as representative of the retail investor population in aggregate.

A central feature is the publicly available *Heat Rank*, which quantifies investor attention from user searching and browsing activity. The platform publishes three related indices: *Searching*, which counts stock searches; *Viewing*, which counts page views; and a composite *Attention* index combining both

[5] A-shares represent domestically listed common stocks denominated in CNY and traded primarily by Chinese residents. B-shares are traded in USD or HKD, were designed for foreign investors, and are substantially less liquid.

signals. Both components are scaled by total activity across all stocks on each day, and stocks are ranked daily by their attention level to produce the real-time *Heat Rank*. Section 3.3 defines these measures formally.

RoyalFlush preserves the *identity* of viewers across time, a feature that, to our knowledge, no other large-scale attention dataset used in finance research provides. Large-scale attention datasets used in prior work (Google Trends (Da et al., 2011), Bloomberg terminal queries (Ben-Rephael et al., 2017), and watchlist additions (Levy et al., 2024)) collapse the viewer set to a scalar aggregate count, discarding who is watching. RoyalFlush records *which* users viewed each stock on each day, enabling a direct, model-free decomposition of aggregate attention into sticky and transitory components. Section 3.4 formalizes this decomposition.

3.2 Complementary Data Sources

Stock prices and firm-level characteristics come from CSMAR (China Stock Market and Accounting Research), the Chinese equivalent of CRSP and Compustat, widely used in academic and institutional research on Chinese equity markets (e.g., Jiang et al., 2022; Chen et al., 2023; Tan et al., 2023).

Firm-specific news is from *SuperMind*, an open-source database aggregating reports from major online financial media outlets. *SuperMind* classifies each article by sentiment (positive, negative, neutral) and allows a single article to be linked to multiple firms, similar in structure to *RavenPack*. We use news counts to test whether attention stickiness effects are driven by news diffusion.

Institutional ownership data come from *iFind*, a commercial database developed by RoyalFlush as a subsidiary platform. *iFind* reports the fraction of shares held by institutions for each A-share firm at quarterly frequency, with comprehensive disclosure in half-year (Q2) and annual (Q4) reports. We use institutional ownership as a cross-sectional moderator to test whether stickiness effects concentrate where retail investors are most likely to be the marginal traders. Appendix Table A.1 defines all variables.

3.3 Attention Measure

The raw attention index, $Attention_{i,t}$, is RoyalFlush's composite daily measure of investor interest in stock i. It combines two activity signals (the number of user searches for the stock, *Searching*, and the number of stock-page views, *Viewing*) with greater weight assigned to views, which dominate in raw counts.[6] Both signals are scaled by total platform activity across all listed stocks on day t, so

[6] The weighting scheme is proprietary but reflects the relative frequency of viewing versus searching activity reported by the platform.

Attention$_{i,t}$ measures stock i's *share* of total daily retail attention on the platform. A value of 0.001 means that 0.1% of all user activity on that day was directed at stock i.

Because the raw index reflects a stock's relative standing within a given day's activity, it is not directly comparable across days with different total platform usage. We therefore convert it into a standardized abnormal measure. Following Da et al. (2011), we define $ATT_{i,t}$ as the log deviation of the daily attention index from its own 30-trading-day rolling median:

$$ATT_{i,t} = \frac{\log(\textit{Attention}_{i,t}) - \widetilde{\text{med}}_{d\in[t-30,\,t-1]}[\log(\textit{Attention}_{i,d})]}{\sigma_i}, \tag{1}$$

where σ_i is the time-series standard deviation of $\log(\textit{Attention}_{i,t})$ for stock i. The rolling median removes the persistent component of each stock's attention level, so $ATT_{i,t}$ captures attention shocks relative to the stock's own recent baseline rather than cross-sectional differences in average coverage. Dividing by σ_i ensures comparability across stocks with different baseline attention volatility. A value of $ATT_{i,t} = 1$ means today's log-attention is one standard deviation above the stock's recent median.

3.4 Attention Stickiness Measure

Building on the individual viewer records described above, we define attention stickiness for stock i on day t at horizon k as:

$$\textit{STK}_{i,t,k} = \frac{\textit{Users View}_{i,\,t-k\,\cap\,t}}{\textit{Users View}_{i,\,t-k}}, \tag{2}$$

where the numerator counts users who viewed stock i's page on both day $t-k$ and day t, and the denominator counts all users who viewed it on day $t-k$, for $k \in \{1, 7, 30\}$ trading days. $\textit{STK}_{i,t,k}$ is therefore the fraction of the day-$t-k$ viewer base that returns to view the same stock k days later. A value of 0.3 at $k = 1$ means 30% of yesterday's viewers revisit today; a value of 0 means the viewer set has turned over entirely. As a measurement, STK records viewer-base overlap: it is a property of the browsing population, distinct from the aggregate attention level (which counts total views regardless of who is watching) and from the autocorrelation of the aggregate attention time series (which captures persistence in the scalar count, not in the viewer identity). STK also does not observe trading volume or order flow. The interpretations under which viewer-base persistence may matter for prices are developed in Section 2.

Figure 2 illustrates the measure. When no prior-day viewer returns (Panel a), $STK = 0$; when one of two returns (Panel b), $STK = 0.5$. Total attention is identical in both cases (two views per day);

only the composition of the viewer set differs.

FIGURE 2 ABOUT HERE

3.5 Sample and Summary Statistics

Our sample covers daily observations for 3,144 A-share firms listed on the Shanghai and Shenzhen Stock Exchanges from January 2018 to December 2024. We restrict the sample to *Main Board* stocks, the core of the listed universe, and exclude firms on the *ChiNext* (Shenzhen) and *STAR* (Shanghai) boards, which have distinct regulatory regimes and firm characteristics. We further exclude firms designated as *ST* or *ST** (Special Treatment, typically imposed after consecutive losses or financial irregularities) and firms with negative shareholders' equity, following Hu et al. (2021), Chen et al. (2023), and Tan et al. (2023).[7] Finally, we exclude the first five trading days following each stock's IPO and the first trading day following any trading suspension and resumption (re-listing). In both cases, the standard $\pm 10\%$ daily price limit does not apply, generating mechanically extreme returns that would distort our return measures.

Table 1 reports summary statistics for the key variables. The raw attention share $Attention_{i,t}$ is highly right-skewed: the average stock accounts for just 0.026% of total daily platform activity, but the maximum reaches 5.02%, meaning that at peak episodes a single stock absorbs more than 5% of all retail investor views. This concentration reflects the well-documented salience effect: a small number of stocks capture a disproportionate share of retail attention on any given day (Barber and Odean, 2008).

The standardized abnormal attention index $ATT_{i,t}$ has a cross-stock mean of 0.079 and standard deviation of 0.655, with a range from -4.06 to 7.54. High-attention episodes (the top decile of the cross-sectional $ATT_{i,t}$ distribution on a given day) correspond to genuine spikes above each stock's recent baseline, not merely periods of persistently high coverage.

Attention stickiness is substantial and highly variable. On average, 38.6% of users who viewed a stock on day $t-1$ return the following day, meaning that on a typical trading day more than one-third of the prior-day viewer base returns to monitor the same stock. The 7- and 30-day stickiness ratios average 33.1% and 23.3%, respectively, confirming that persistent monitoring extends well beyond adjacent days. The maximum 1-day stickiness ratio approaches 87% (for some firms, nearly the entire prior-day viewer base returns), while cross-stock variation is large. This combination of high average

[7] ST and ST*-designated stocks face tighter daily price limits of $\pm 5\%$ versus the standard $\pm 10\%$.

values and wide dispersion makes STK empirically tractable as both a level variable and a classification instrument.[8]

TABLE 1 ABOUT HERE

The left panel of Figure 3 illustrates these dynamics for Shanghai Fosun Pharmaceutical (600196). The attention index (upper panel) spikes sharply in tandem with news bursts (lower panel), while 1-day attention stickiness (middle panel) is substantially smoother and less news-reactive. This visual decorrelation anticipates a central empirical pattern: unlike aggregate attention, stickiness does not simply track information arrival and varies largely independently of news flow.

FIGURE 3 ABOUT HERE

3.6 What Drives Attention vs. Stickiness?

To confirm that the two measures capture distinct dimensions of investor behavior, we regress each on contemporaneous firm characteristics, trading variables, and news indicators, controlling for the lagged dependent variable. Both dependent variables are normalized, making coefficient magnitudes directly comparable across panels. Newey-West standard errors with the optimal bandwidth address serial correlation concerns. Columns (1)–(4) omit fixed effects; column (5) includes firm and time fixed effects.

Panel A of Table 2 reveals that normalized abnormal attention is highly persistent: the lagged dependent variable coefficient exceeds 0.70 across all specifications, and the simplest model with contemporaneous returns and lagged attention alone achieves an adjusted R^2 of 0.768. Contemporaneous returns affect attention asymmetrically: positive returns are roughly five times more powerful than negative returns (24.4 vs. -5.3 in column (1)), consistent with retail investors being drawn disproportionately to rising stocks. News indicators, turnover, and firm fundamentals add modest incremental explanatory power; the full specification with firm and time fixed effects reaches an adjusted R^2 of 0.789. These patterns are consistent with prior literature (Da et al., 2011; Levy et al., 2024).

Panel B reveals a stark contrast in dynamics. Attention stickiness exhibits weak mean reversion rather than persistence: the lagged STK coefficient is small and negative (-0.02 to -0.06), implying

[8] Correlations among key variables are reported in Appendix Table A.2.

that high stickiness today slightly predicts lower stickiness tomorrow. Contemporaneous positive returns are the strongest single driver (28.2 in column (1)), with a response to negative returns that is much weaker (-3.2) and shrinks further once trading variables are included. News indicators, which raise attention in Panel A, become insignificant or negative for stickiness once fundamentals and fixed effects are controlled (column (5)), consistent with news episodes bringing in new viewers who dilute the returning viewer base. Even the full specification explains only 31.9% of stickiness variation, leaving roughly two-thirds of day-to-day variation in viewer-base composition unexplained by observables.

The cross-panel comparison shows that stickiness and attention have fundamentally different dynamics despite responding to similar contemporaneous forces. The persistence gap (lagged coefficient of 0.71 for attention vs. -0.05 for stickiness) means that attention shocks propagate forward while stickiness shocks do not; the explanatory-power gap (adjusted R^2 of 78.9% vs. 31.9%) means that most day-to-day variation in stickiness is left unexplained by observable firm characteristics and news flow. A stock can attract high aggregate attention while retaining few of the previous day's viewers (low STK), or attract modest attention from a largely returning base (high STK). Aggregate attention proxies conflate these two situations by collapsing the viewer set to a scalar count; our measure separates them.

TABLE 2 ABOUT HERE

4 Analysis of Sticky and Transitory Attention

We first establish that abnormal retail attention predicts short-term overpricing followed by rapid reversal in the Chinese market, replicating the pattern documented for U.S. markets. We then decompose the attention-return relationship using our baseline continuous stickiness specification and a companion indicator specification, showing that a one-standard-deviation increase in stickiness is associated with roughly twice the same-day return of an attention shock of the same magnitude. Under the binary classification, conditional on the same level of abnormal attention, the estimated same-day response is over 112 basis points when retention of the previous day's viewers is above the panel median and essentially zero when it is below. The level of stickiness also contributes independently, predicting lower future returns beyond its role as a moderator of the attention shock. We conclude with an orthogonalization test confirming that these effects are not driven by co-movement with observables.

4.1 Revisiting the Role of Attention in the Chinese Market

Prior studies document that attention-grabbing stocks experience short-term overpricing followed by reversal (e.g., Da et al., 2011; Drake et al., 2017; Levy et al., 2024; Da et al., 2025). We establish that this pattern holds in our Chinese sample using the standardized abnormal attention index $ATT_{i,t}$ defined in Section 3. Panel (a) of Figure 4 plots average returns around high-attention episodes (top-decile abnormal attention, non-overlapping 21-day windows): stocks exhibit a sharp return of approximately 3% on the event day, followed by rapid reversal over subsequent trading days.[9]

We estimate the return dynamics formally using

$$r_{i,t\to t+h} = \beta_h \times ATT_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}, \tag{3}$$

where the dependent variable is the cumulative holding-period return from the closing price on day $t-1$ to the closing price on day $t+h$. Firm fixed effects absorb permanent cross-sectional differences in average returns and attention levels; time fixed effects absorb market-wide shocks common to all stocks on a given day. The identifying variation is therefore within-stock, day-to-day: how does a given stock's return respond when its own attention rises above its recent baseline? Standard errors are clustered by firm and time.

TABLE 3 ABOUT HERE

Panel (a) of Figure 5 plots the expected return dynamics following a one-standard-deviation shock to abnormal attention. The expected return peaks at over 80 bps on the event day and decays steadily over subsequent horizons, remaining positive through day 40 as the initial price impact gradually fades. The marginal attention coefficient nonetheless turns negative at forward horizons: shifting the return window to $r_{i,t+1\to t+h}$, Panel A of Table 3 quantifies the reversal: -4.63 bps at $h=1$, deepening to -25.8 bps at $h=5$, -39.8 bps at $h=10$, and -69.9 bps at $h=40$.[10] The 40-day reversal is roughly three times the 22 bps documented by Da et al. (2025) for the U.S. market, consistent with faster reversals in China's retail-dominated setting (see Da et al., 2014, for the anatomy of short-horizon reversals).

[9] The pattern is almost identical when reporting market-adjusted returns.

[10] Panel B replaces $ATT_{i,t}$ with a high-attention indicator, confirming the pattern: -6.40 bps at $h=1$, -25.3 bps at $h=5$, and -37.0 bps at $h=10$ (Panel B of Table 3).

FIGURES 4 AND 5 ABOUT HERE

4.2 Main results

The baseline regression treats all attention shocks symmetrically. We now ask whether the price response differs depending on who is watching. The test continues to exploit within-stock variation: firm and time fixed effects ensure that the estimates reflect how the *same stock's* return responds differently when its viewer-base retention is high versus low. Panel (b) of Figure 4 previews the answer: stocks experiencing sticky attention display a substantially larger same-day return and a more pronounced reversal than stocks with transitory attention, where the immediate price impact is negligible.

The regression decomposes the attention-return relationship into three components. We augment the baseline specification as follows:

$$r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}, \quad (4)$$

where $STK_{i,t}$ is the normalized continuous attention stickiness measure and all other variables are defined as before. Each coefficient answers a different question. $\beta_{1,h}$ measures the marginal effect of a one-standard-deviation attention shock evaluated at the unconditional mean of stickiness. $\beta_{2,h}$ captures the marginal effect of a one-standard-deviation increase in stickiness evaluated at the mean of attention — a direct test of whether the composition of the viewer base predicts returns independently of the attention level (Prediction P3). $\beta_{3,h}$ measures how much a one-standard-deviation increase in stickiness amplifies or attenuates the attention-shock effect.

The continuous specification is our baseline because it preserves the full cross-sectional variation in stickiness intensity rather than collapsing it to a binary state. Panel A of Table 4 reports estimates from this specification.[11] As a complement, Panel B replaces $STK_{i,t}$ with an indicator $\mathbb{I}_{i,t}^{STK}$ equal to one when stickiness exceeds the panel median. The indicator specification sacrifices intensity information but yields a transparent decomposition into sticky and transitory states; all return-dynamics figures except Figure 6 use this indicator-based specification for visual clarity. Tightening the cutoff to the top and bottom deciles widens the sticky–transitory asymmetry (Online Appendix Figure A.5).

[11] We replicate the test using alternative settings to produce abnormal attention in Online Appendix Table A.7. The abnormal attention is calculated as the log deviation of the daily attention index from its own 8-week, 4-week, and 2-week medians, respectively.

4.2.1 Contemporaneous price impact

All three coefficients of the continuous-STK specification are large and highly significant at $h = 0$ (Panel A of Table 4). A one-standard-deviation attention shock, evaluated at mean stickiness, is associated with a same-day return of $\hat{\beta}_1 = 31.03$ bps. A one-standard-deviation increase in stickiness, evaluated at mean attention, contributes $\hat{\beta}_2 = 73.93$ bps — roughly twice the marginal attention effect, establishing stickiness intensity as the dominant contemporaneous correlate of returns. The interaction $\hat{\beta}_3 = 40.90$ bps is positive, confirming that high stickiness amplifies the attention shock. The total same-day return from a joint one-standard-deviation shock to both ATT and STK is $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 = 145.9$ bps.

The indicator specification in Panel B translates these continuous gradients into a binary contrast. Panel (b) of Figure 5 plots expected return dynamics separately for the sticky and transitory states. The sticky-attention line spikes to over 220 bps on the event day while the transitory-attention line remains near zero. The underlying regression coefficients sharpen the decomposition: the pure attention-shock effect in the sticky state is $\hat{\beta}_1 + \hat{\beta}_3 = 112.57$ bps, while the same shock in the transitory state yields $\hat{\beta}_1 = 0.91$ bps, essentially zero. The gap is the paper's core finding: conditional on the same level of abnormal attention, the estimated response is large and positive when viewer-base retention is high, and essentially zero when retention is low, consistent with Prediction P1. The contrast is not one of degree but of kind: sticky attention is associated with overreaction comparable in magnitude to earnings-announcement returns, while transitory attention is associated with none.

The indicator specification also reveals a separate level effect: $\hat{\beta}_2 = 107.58$ bps at $h = 0$, consistent with a stable viewer base being associated with persistent buying pressure throughout the day. The total same-day return for a sticky-attention stock with a one-standard-deviation ATT shock is $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 = 220$ bps — a total directly visible as the sticky-attention expected return at $h = 0$ in Panel (b) of Figure 5. This same-day gain fades steadily over subsequent horizons: the expected return for sticky-attention stocks declines toward zero by day 40, consistent with transient demand rather than permanent information (Prediction P2). The sign of $\hat{\beta}_2$ itself reverses at forward horizons (turning negative by $h = 1$), as the within-day buying pressure dissipates and prices correct.

Figure 6 visualizes the continuous-specification dynamics (Panel A) across the full horizon. The direct effect of attention stickiness ($\hat{\beta}_{2,h}$, orange) dominates the direct effect of abnormal attention ($\hat{\beta}_{1,h}$, blue) at every horizon: stickiness contributes roughly 75 bps on the event day compared with roughly 30 bps for attention, and drives the bulk of the subsequent decline. The total expected return from a joint one-standard-deviation shock to both ATT and STK ($\hat{\beta}_{1,h} + \hat{\beta}_{2,h} + \hat{\beta}_{3,h}$, green) peaks at

approximately 145 bps and fades toward zero by day 40, with a positive interaction ($\hat{\beta}_{3,h} > 0$ at short horizons) confirming that high stickiness amplifies the attention shock. These patterns are quantitatively consistent with the indicator-based results in Panel B: the sticky–transitory asymmetry does not depend on whether stickiness is measured as a continuous intensity or a binary state.

FIGURE 6 ABOUT HERE

4.2.2 Return predictability: sticky–transitory asymmetry

We now examine whether the two specifications predict different post-event return dynamics. Replacing the dependent variable with $r_{i,t+1\to t+h}$ shifts the return window forward.

In the continuous specification (Panel A), the marginal attention effect $\hat{\beta}_1$ turns negative immediately (-8.52 bps at $h = 1$) and deepens monotonically to -39.40 bps at $h = 40$. The interaction $\hat{\beta}_3$ remains positive and significant throughout ($+7.17$ bps at $h = 1$), indicating that high stickiness delays the reversal, but stays an order of magnitude smaller than the direct stickiness effect at every forward horizon. The direct stickiness effect $\hat{\beta}_2$ is the dominant forward predictor: -1.92 bps at $h = 1$, deepening to -18.21 bps at $h = 5$, -33.84 bps at $h = 10$, -62.68 bps at $h = 20$, and -112.62 bps at $h = 40$, with no sign of flattening at the boundary.

The indicator specification (Panel B) delivers a qualitatively identical pattern, visible in the two expected return paths in Panel (b) of Figure 5. For transitory-attention stocks, abnormal attention predicts an immediate decline of 11.41 bps on the following day ($\hat{\beta}_1$), while sticky-attention stocks show a sharply attenuated day+1 response ($\hat{\beta}_3 = +10.90$ bps), leaving the net ATT-driven effect near zero on day $t+1$. The timing asymmetry is consistent with the demand-concentration interpretation: a stable viewer base is associated with more persistent price pressure, while a rotating crowd of newcomers is associated with faster correction. By $h = 10$, the interaction coefficient $\hat{\beta}_3 = -2.51$ bps is statistically indistinguishable from zero, yet the two expected return paths remain well separated throughout the 40-day window: the persistent gap reflects the ongoing contribution of the direct stickiness level effect ($\hat{\beta}_2$), which grows in magnitude at longer horizons. Both specifications therefore agree: the interaction operates only on impact; the dominant predictor of multi-week expected returns is the level of stickiness, not the differential response to the attention shock (Prediction P3).

TABLE 4 ABOUT HERE

4.2.3 Direct predictability of attention stickiness

Does stickiness intensity predict returns independently of the attention shock? The coefficient $\hat{\beta}_{2,h}$ answers this question: in the continuous specification (Panel A), it measures the return effect of a one-standard-deviation increase in stickiness at horizon h, holding the attention level at its mean.

The answer is unambiguous. The $\hat{\beta}_2$ magnitudes reported above — deepening from -1.92 bps at $h = 1$ to -112.62 bps at $h = 40$ in the continuous specification, and reaching -160.44 bps in the indicator specification — far exceed the interaction-based differential ($\hat{\beta}_{3,h}$) at every horizon beyond five days. The stickiness level itself, not its interaction with the attention shock, is the dominant forward-return predictor at multi-week horizons. This predictive channel is absent from aggregate attention proxies by construction, since they collapse the viewer set to a scalar count.

A natural question is whether STK is endogenous to contemporaneous returns: do investors revisit precisely because prices rose, making stickiness a consequence rather than a cause of overpricing? Table 2 confirms that contemporaneous positive returns are indeed the strongest single driver of stickiness, so the concern has empirical support. Two features of the data nonetheless argue against a return-chasing interpretation of our predictive results. First, lagged STK enters negatively (-0.05), implying that high stickiness today slightly predicts *lower* stickiness tomorrow rather than the persistence that a return-chasing habit would generate. Second, and more importantly, STK predicts *lower* future returns: return-chasing behavior implies positive continuation, not the sharp reversals we document. This sign structure also addresses the concern that persistent time-varying firm characteristics (e.g., the composition of a stock's retail following) might jointly drive high STK and lower returns. Such a mechanism would need to explain why stickiness predicts lower returns only at forward horizons while generating a positive same-day premium ($\hat{\beta}_2 > 0$ at $h = 0$), a pattern more naturally consistent with transient demand that initially elevates prices and subsequently corrects. The alternative attention specification in Section 4.3 and an orthogonalized measure that removes the component of STK linearly related to attention levels and firm characteristics (Figure 8) indicate that the results are not driven by co-movement with included observables.

A broader interpretive point deserves emphasis. Our data observe browsing, not trading: we measure which investors view a stock's page, not which investors submit orders. The interpretation that viewer-base persistence translates into concentrated buying pressure is therefore indirect. Four pieces of evidence support this inference: (i) the sharp same-day return asymmetry between sticky and transitory attention, which is difficult to explain without some form of differential trading behavior; and (ii) the full

subsequent reversal, consistent with transient demand rather than permanent information. Two further pieces come from outside the panel regressions: (iii) the monotone relationship between STK and turnover documented in Section 5, linking viewer-base persistence to trading activity at the portfolio level; and (iv) the concentration of effects in retail-dominated market segments (Section 6), where browsing on a retail platform is most likely to precede trading. Individual-level data linking browsing records to brokerage accounts would provide a more direct test of this mechanism.

The STK level effect ($\beta_{2,h} = -112.62$ bps over forty trading days in the continuous specification, -160.44 bps in the indicator specification) and the Low-Low portfolio outperformance documented in Section 5 (1.311% per month) provide complementary evidence from different sources of variation. The panel regressions isolate within-stock variation through firm fixed effects, identifying how a given stock's return changes when its stickiness state shifts. The portfolio sorts exploit between-stock variation in the cross-sectional level of retail monitoring, identifying which stocks systematically earn a neglect discount. Both results are consistent with attention composition mattering for prices, but they are identified from different variation and need not reflect the same underlying mechanism.

4.3 Robustness: Orthogonalized Stickiness

A central concern is that STK may proxy for aggregate attention levels, firm characteristics, or news flow rather than capturing a genuinely distinct dimension of investor behavior.

We construct $STK_{i,t}^{\perp}$ by regressing STK on contemporaneous and lagged attention, lagged return, log market capitalization, turnover, and indicators for positive, negative, and neutral news, then retaining the intercept plus residual (Appendix Table A.3). By construction, $STK^{\perp}$ is orthogonal to all included controls. Because these controls include the lagged but not the contemporaneous return, $STK^{\perp}$ does not purge the component of stickiness that co-moves with the same-day return. The orthogonalization therefore addresses confounding by attention level, size, liquidity, and news, but not the mechanical association between same-day returns and revisiting, and we accordingly read the same-day estimates as associations.

Re-estimating specification (4) with $\mathbb{I}_{i,t}^{STK}$ replaced by the orthogonalized indicator leaves the contemporaneous sticky-attention effect ($\hat{\beta}_1 + \hat{\beta}_3$) essentially unchanged from the Panel B estimate of 112.57 bps. Panel (a) of Figure 8 plots expected return dynamics under the orthogonalized measure: the sticky-attention expected return path (triangle-dotted) closely tracks the baseline, confirming that the sharp same-day spike and subsequent fade survive the removal of shared variation with observables.

The result mitigates the concern that STK is a relabeling of aggregate attention or a mechanical byproduct of firm size, liquidity, or news: the sticky–transitory asymmetry survives after removing all variation shared with these observables. This does not resolve unobserved heterogeneity, but it narrows the set of plausible confounds to time-varying firm characteristics that are both orthogonal to the included controls and capable of producing the specific sign pattern documented above.

5 Portfolio Level Analysis

The panel regressions in Section 4 identify how attention shocks affect prices *within* a stock's own time series, but by design absorb all cross-sectional level differences in average attention and stickiness across stocks. This section asks a distinct question: does the *cross-sectional level* of retail monitoring a stock receives (not a transient spike above its own history, but its standing relative to other stocks) predict future returns? The source of variation is different: sorting on prior-month average ATT and STK exploits between-stock differences in monitoring intensity. The portfolio results are not mechanically implied by the panel evidence, because the panel regressions condition on firm fixed effects and therefore cannot speak to whether stocks with persistently low (or high) attention and stickiness earn different average returns. The goal of the portfolio analysis is to validate the cross-sectional implications of attention composition, not to propose an implementable trading rule. All reported returns are gross of transaction costs, and the short leg involves stocks that may be costly or difficult to sell short.

To establish a baseline, we form value-weighted quintile portfolios on prior-month average abnormal attention, rebalanced monthly. Panels (a) and (b) of Figure 7 show that attention alone already matters: the low-ATT quintile consistently outperforms the high-ATT quintile, and the long-short spread remains positive throughout the sample. The ATT-only sort, however, masks substantial heterogeneity within quintiles: two stocks with the same average ATT can differ sharply in how many of the previous day's viewers return. Panels (c) and (d) show what happens when stickiness is added. Panel (c) plots cumulative returns for the four corner portfolios of a 5×5 double-sort: the Low-Low portfolio (low ATT, low STK) separates upward from all others, while the High-High portfolio (high ATT, high STK) declines steadily. The off-diagonal corners (high ATT with low STK, or low ATT with high STK) fall in between, confirming that neither dimension alone produces the extreme outcomes. Panel (d) plots the Low-Low minus High-High spread, which remains positive for nearly the entire sample. The economic takeaway from the four panels is that attention level matters, but attention composition sharpens the cross-sectional pattern substantially.

FIGURE 7 ABOUT HERE

Returns decrease monotonically in both dimensions of the 5×5 grid of value-weighted portfolios sorted jointly on ATT and STK, whose average ex-post monthly return and turnover appear in Panels A and B of Table 5. Two patterns in Panel A define the central result. First, the Low-Low corner (ATT-Q1, STK-Q1) earns 1.311% per month; the High-High corner (ATT-Q5, STK-Q5) earns -0.922%, a spread of 2.23 percentage points per month. Second, the two off-diagonal corners are substantially milder: Q5/Q1 (high-ATT, low-STK) earns 0.441% and Q1/Q5 (low-ATT, high-STK) earns 0.206%. The extreme return outcomes require *both* dimensions simultaneously; neither sort alone produces the full spread.

The STK gradient across the 5×5 grid exceeds the ATT gradient. Averaging across all ATT quintiles, the pure STK spread is 1.02 percentage points per month (STK-Q1 average: 0.876%; STK-Q5 average: -0.147%). The analogous pure ATT spread, averaging across STK quintiles, is 0.78 percentage points per month (ATT-Q1 average: 0.795%; ATT-Q5 average: 0.016%). The STK gradient thus exceeds the ATT gradient by 24 basis points per month, the cross-sectional analog of the finding in Section 4 that $\beta_{2,h}$ dominates $\beta_{3,h}$ at longer horizons. Within the highest-attention quintile (ATT-Q5), stickiness alone creates a 1.36 percentage point spread: Q5/Q1 stocks earn 0.441% per month while Q5/Q5 stocks earn -0.922%. Even among stocks receiving the most aggregate retail attention, the return outcome depends on whether the same investors return.

TABLE 5 ABOUT HERE

Turnover mirrors the return pattern (Panel B of Table 5). The High-High corner trades at 4.1 times the volume of the Low-Low corner (monthly turnover 0.545 vs. 0.133), and the gradient is monotone along both dimensions: within ATT-Q5, turnover rises from 0.178 at STK-Q1 to 0.545 at STK-Q5. This joint return-turnover pattern supports the demand-concentration interpretation. For Low-Low stocks, sparse trading points to persistent underpricing: with too few transactions to aggregate dispersed information, prices adjust slowly over months. For High-High stocks, elevated turnover reflects the correlated monitoring associated with same-day overreaction in Section 4. Although we observe browsing rather than individual trades, the monotone relationship between STK and turnover across the grid supports the interpretation that repeated monitoring is associated with more concentrated trading activity.

The outperformance of the Low-Low portfolio is consistent with a neglect discount rather than a reversal premium. Under the investor recognition hypothesis of Merton (1987), stocks trade at a persistent discount when the investor base is too narrow to fully incorporate their value. Stocks doubly off the retail radar (low-ATT and low-STK) face the deepest recognition gap: even the few investors who do visit almost never return, so no stable viewer base develops that could translate into the buying pressure needed to close the discount. As coverage gradually expands, the correction accumulates over months rather than days. The mechanism on the short leg is different: High-High stocks display the overreaction-and-reversal pattern documented in Section 4, consistent with concentrated demand from a persistent viewer base temporarily elevating prices before the correction completes within approximately forty trading days. The spread grows monotonically: 2.24%, 4.45%, 6.62%, 8.93%, 11.35%, and 13.74% at one through six months, as the long leg's slow neglect correction compounds the short leg's faster behavioral reversal.

TABLE 6 ABOUT HERE

The cross-sectional spread is economically large: annualized Sharpe ratios reach 1.375 at the one-month horizon and 1.648 at six months (Panel A of Table 6). The long leg alone (the Low-Low portfolio, earning 1.311% per month) outperforms the value-weighted market without requiring short-selling. Panel B examines the risk structure. The long-short spread carries a strongly negative market beta ($\hat{\beta}^{MKT} = -34.0$, significant at the 1% level), reflecting the defensive character of neglected stocks on the long side and the high-beta, high-sentiment character of the short side. The CAPM alpha is 2.30% per month (significant at the 1% level). Adding the sentiment factor PMO (loading $+87.1$, significant at the 1% level) brings the four-factor alpha to 1.455% per month (significant at the 1% level) with an R^2 of 0.39. The large PMO loading indicates that High-High stocks sit at the peak of the retail sentiment cycle; the residual alpha of 1.455% represents a return spread that standard risk models do not capture.[12] The spread is unlikely to reflect a pure size or illiquidity premium: the four-factor model already includes SMB, and excluding the least liquid stocks leaves the alpha essentially unchanged at 1.466% per month (last column of Table 6).

These magnitudes should be interpreted as indicative of the economic importance of the cross-sectional pattern rather than as realizable trading profits or precise point estimates; estimated over

[12]Pricing factors include market excess return, SMB (size effect), HML (value effect), and PMO (sentiment effect). Factors are constructed following Liu et al. (2019); historical data are downloaded from Yuan's website.

a seven-year sample, the Sharpe ratios and alphas carry wide standard errors. The reported returns are gross of transaction costs, and the High-High short leg (turnover 0.545) would face substantial implementation frictions in practice.

The portfolio evidence and the panel regressions address different questions using different sources of variation. The panel results show that within-stock shifts in attention composition predict short-horizon overreaction and reversal. The portfolio results show that cross-sectional differences in average monitoring levels are associated with longer-run return spreads, separating genuinely neglected stocks from temporarily quiet ones and persistently overmonitored stocks from briefly salient ones. Both findings are consistent with attention composition mattering for prices, but they operate at different frequencies and through different mechanisms.

6 Mechanism Tests and Robustness

Prediction P4 states that the high- versus low-retention reversal gap should be larger where retail investors are more likely to be the marginal traders and arbitrage is more constrained. We test this using two proxies for retail participation (nominal share price and institutional ownership), then aggregate attention and stickiness to equally-weighted industries to confirm the findings are not confined to small or illiquid stocks. Additional robustness checks are collected in Online Appendix B; volatility dynamics appear in Online Appendix C.

6.1 Nominal Price and Institutional Ownership

In the Chinese A-share market, trades are subject to the board-lot rule, which requires a minimum trading unit of 100 shares. Because the cash outlay per lot scales mechanically with the nominal share price, high-priced stocks are less accessible to small, capital-constrained retail investors, whereas low-priced stocks are more affordable and thus more likely to be retail-dominated. Du et al. (2025) document the absence of mid-term momentum among low-price stocks, consistent with this retail clientele effect.

We sort stocks into quintiles by prior-month closing price and rebalance monthly, re-estimating the return prediction specification within the lowest and highest quintiles. The stickiness-driven reversal concentrates among low-price stocks (Panel A of Table 7): the stickiness level predicts -115.64 bps over forty trading days in the low-price quintile versus -80.00 bps in the high-price quintile, and the High-minus-Low difference of 35.64 bps is significant at the 1% level. The interaction term shows the same asymmetry: at $h = 10$ it is -0.44 bps and statistically indistinguishable from zero for low-price

stocks but +5.45 bps (significant at the 1% level) for high-price stocks, a High-minus-Low gap of 5.89 bps (significant at the 5% level). Where share prices impose a natural barrier on small-account participation, the stickiness channel weakens, consistent with capital-constrained retail investors being the marginal traders in the low-price segment.

TABLE 7 ABOUT HERE

We complement this affordability-based proxy with a more direct measure of institutional ownership. We collect institutional holdings from *iFind* (see Section 3), which reports the fraction of shares held by institutions for each A-share stock at the quarterly frequency. We sort stocks into quintiles based on Q4 institutional ownership and rebalance annually.[13] Attention stickiness-induced price pressure and reversal both concentrate in stocks with low institutional ownership (Panel B of Table 7). On the event day, the direct stickiness effect is 79.09 bps in the lowest ownership quintile versus 66.31 bps in the highest, and the amplifying interaction is 43.50 versus 35.07 bps; both High-minus-Low differences (-12.78 and -8.43 bps) are significant at the 1% level. The subsequent reversal is correspondingly deeper: the stickiness level predicts -118.16 bps over forty trading days in the low-ownership quintile versus -91.22 bps in the high-ownership quintile, a difference of 26.94 bps (significant at the 5% level). This pattern is consistent with institutional presence dampening transient price pressure: where informed arbitrageurs can trade freely, the price movements associated with high viewer retention are smaller and leave less to reverse. Both results support prediction P4: the stickiness effects concentrate where retail investors are likely to be the marginal traders and institutional arbitrage is limited.

The nominal price and institutional ownership results are difficult to reconcile with a simple rational-learning interpretation: under rational learning, repeated monitoring improves pricing accuracy and should not concentrate in stocks where informed institutions are absent. The cross-sectional patterns instead favor behavioral reinforcement as the preferred interpretation.

6.2 Industry-level Predictability

Stock-level results could be driven by small or illiquid firms whose idiosyncratic reversals cancel in aggregate. To address this concern, we aggregate firm-level attention and stickiness to equally-weighted industries following Da et al. (2025) and re-estimate our baseline specification at the industry level.

[13] Chinese disclosure rules require comprehensive reporting only in half-year (Q2) and annual (Q4) reports; Q1 and Q3 filings typically cover only top positions. Using Q4 ownership avoids the measurement error from incomplete disclosure. Tests with quarterly sorting generate similar results.

For each trading day and industry, we weight firm-level variables equally across constituent firms.[14] We also apply the same scheme to form equally-weighted industry returns. The right panel of Figure 3 plots the time-series patterns of equally-weighted attention and attention stickiness, and total news counts of the chemical fiber manufacturing industry. Consistent with stock-level patterns, aggregate attention co-moves closely with news intensity, while attention stickiness is substantially less responsive to news flow, consistent with the low explanatory power of news for stickiness documented in Table 2.

We then re-estimate specification 4 with all variables converted to the industry level. Panel (b) of Figure 8 reports the expected return dynamics at the industry level. The key visual pattern mirrors the stock-level evidence in Panel (b) of Figure 5: the sticky-attention expected return (triangle-dotted) spikes at $h = 0$ while the transitory-attention path (square-dotted) remains near zero. The gap between the two paths narrows gradually over subsequent horizons, with the sticky-attention expected return decaying from its same-day peak but remaining positive through day 40. Transitory-attention industries show a muted same-day response and drift slightly below zero at longer horizons, consistent with diffuse demand that generates little aggregate price pressure.

The industry-level results indicate that sticky and transitory attention are associated with systematic, portfolio-relevant pricing patterns rather than idiosyncratic noise that cancels in the cross-section. The large same-day gap between the two expected return paths and the gradual decay of the sticky-attention expected return at the industry level mirror the stock-level dynamics, suggesting that the sticky–transitory asymmetry operates at a scale relevant for diversified portfolios.

FIGURE 8 ABOUT HERE

6.3 Additional Robustness Checks

We subject the baseline results to seven robustness tests, each targeting a distinct identification concern: (i) event contamination by quarterly disclosure windows, (ii) sensitivity to the choice of attention proxy, (iii) sensitivity to the horizon used to define stickiness, (iv) sensitivity to the stickiness cutoff, (v) cross-sectional heterogeneity in where the effects concentrate, (vi) sensitivity to the functional form of both attention and stickiness, and (vii) sensitivity to the window used to compute abnormal attention. Details and results are presented in Online Appendix B.

[14] Industries follow the 2023 China Securities Regulatory Commission (CSRC) Industry Code.

These checks together establish that the sticky–transitory asymmetry does not depend on the baseline measurement choices. The asymmetry survives excluding quarterly disclosure windows, replacing the composite attention measure with separate searching and viewing indices, extending the stickiness horizon to seven or thirty days, and replacing both continuous measures with above-median indicators; tightening the classification cutoff to cross-sectional deciles widens rather than narrows the gap. The cross-sectional heterogeneity results add an economic dimension: the sticky-versus-transitory gap is largest precisely where arbitrage is most constrained, among short-sale restricted stocks with low institutional participation, consistent with the limits-of-arbitrage interpretation.

A final measurement check addresses the construction of the abnormal-attention index itself. Recomputing $ATT_{i,t}$ against 40-, 20-, and 10-trading-day medians, rather than the 30-day baseline, leaves the contemporaneous stickiness premium, its interaction with attention, and the negative forward predictability of stickiness qualitatively unchanged and quantitatively similar (Online Appendix Table A.7). The central finding, that the temporal composition of attention rather than its level drives short-term price dynamics, is therefore not an artifact of the window used to scale attention shocks.

6.4 Impact of Sticky Attention on Stock Volatility

Return predictability tests the mean response to attention shocks; conditional volatility provides a second-moment test. Rational information-based theories predict a volatility response that is roughly symmetric across sticky and transitory attention (Andrei and Hasler, 2015), while behavioral demand concentration predicts amplification by stickiness. The data support amplification: a one-standard-deviation attention shock raises next-day conditional volatility by 3.71 bps at average stickiness and by 10.28 bps one standard deviation higher, and stickier stocks carry a direct volatility premium of 6.12 bps independent of the attention shock. Online Appendix C presents the GARCH specification, the unconditional response, and the full horizon profile.

7 Conclusion

The temporal composition of retail investor attention predicts whether attention moves prices. Using individual browsing records that identify the same investor across consecutive days, we show that, conditional on the same level of abnormal attention, the estimated same-day response exceeds 112 basis points, fully reversed within forty trading days, when the stock retains an unusually large share of the previous day's viewers, and is essentially zero when retention is unusually low. Aggregate attention

proxies assign the same number to both episodes.

Attention stickiness operates at two frequencies. At daily horizons, the stickiness level itself predicts lower future returns independently of any concurrent attention shock, a channel absent from aggregate attention proxies by construction. At monthly horizons, the cross-sectional level of retail monitoring separates genuinely neglected stocks from temporarily quiet ones, producing a neglect discount of 13.74% over six months against the high-attention, high-stickiness portfolio.

The evidence is consistent with viewer-base persistence being associated with stronger transient price pressure, particularly where retail investors are likely to be the marginal traders. High retention is associated with larger overreaction and fuller reversal, not with more accurate pricing, a pattern difficult to reconcile with a simple rational-learning interpretation; behavioral reinforcement is our preferred interpretation, though the reduced-form design does not identify the mechanism directly. The effects concentrate where the behavioral channel predicts: low-price stocks and low-institutional-ownership firms, where institutional arbitrage is most constrained. Our individual-level data make this comparison possible; aggregate attention proxies cannot distinguish a returning viewer from a new one, so the two channels are observationally equivalent under such measures.

Four limitations deserve acknowledgment. First, our identification relies on reduced-form panel regressions with firm and time fixed effects, not on exogenous variation in attention composition. We cannot rule out that persistent time-varying unobservables jointly influence STK and returns, though the sign structure (positive contemporaneous effect reversing to negative at forward horizons) and the stability of the estimates after removing the component of STK related to observables argue against a simple confounding story. Second, the analysis observes browsing behavior, not trading: the behavioral interpretation requires that repeated monitoring translates into concentrated buying, a link the turnover evidence supports at the portfolio level but that individual-level brokerage data could test directly. Third, the sample period (2018–2024) includes significant market events, including the COVID-19 disruption and episodes of regulatory intervention in Chinese equity markets; although the sticky–transitory asymmetry is present throughout the sample, formal sub-period stability tests would further strengthen the evidence. Fourth, the analysis covers a single market and a single platform. China's retail-dominated structure, board-lot constraints, and short-sale restrictions amplify the behavioral channel, making it more detectable. The underlying concept, however, is portable: attention composition shapes order-flow structure in any market where investors can revisit the same information source. The psychological forces at work (commitment escalation, confirmation bias, familiarity) are

general, and our cross-sectional results (Prediction P4) imply that the effects should be smaller, not absent, in markets with higher institutional participation and fewer trading frictions.

The results point to a dimension of attention that is both measurable and empirically relevant. Platform features that surface the same stocks to the same users repeatedly, such as personalized watchlists, recommendation algorithms, and curated trending lists, may systematically concentrate retail demand in ways that aggregate attention measures cannot detect. The way investors pay attention, not merely how much, is a first-order predictor of short-term price dynamics.

References

Andrei, Daniel and Michael Hasler (2015) "Investor attention and Stock Market Volatility," *Review of Financial Studies*, Vol. 28, No. 1, pp. 33–72, January.

Andrei, Daniel, Henry Friedman, and N. Bugra Ozel (2023) "Economic Uncertainty and Investor Attention," *Journal of Financial Economics*, Vol. 149, No. 2, pp. 179–217.

Barber, Brad M. and Terrance Odean (2008) "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *Review of Financial Studies*, Vol. 21, No. 2, pp. 785–818, April.

Barber, Brad M., Xing Huang, Terrance Odean, and Christopher Schwarz (2022) "Attention-Induced Trading and Returns: Evidence From Robinhood Users," *The Journal of Finance*, Vol. 77, No. 6, pp. 3141–3190, December.

Ben-Rephael, Azi, Bruce I. Carlin, Zhi Da, and Ryan D. Israelsen (2021) "Information consumption and Asset Pricing," *The Journal of Finance*, Vol. 76, No. 1, pp. 357–394, February.

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen (2017) "It Depends on Where You Search: Institutional Investor Attention and Underreaction to News," *Review of Financial Studies*, Vol. 30, No. 9, pp. 3009–3047, September.

Ben-Rephael, Azi, Zhi Da, Peter D. Easton, and Ryan D. Israelsen (2022) "Who pays Attention to SEC Form 8-K?" *The Accounting Review*, Vol. 97, No. 5, pp. 59–88, September.

Chemmanur, Thomas and An Yan (2009) "Product Market Advertising and New Equity Issues," *Journal of Financial Economics*, Vol. 92, No. 1, pp. 40–65, April.

Chen, Xin, Li An, Zhengwei Wang, and Jianfeng Yu (2023) "Attention Spillover in Asset Pricing," *The Journal of Finance*, Vol. 78, No. 6, pp. 3515–3559.

Cookson, J. Anthony, Diego Garcia, and Elvis Jarnecic (2024) "News Consumption in the Wild," *Working Paper, Avaliable at SSRN 4692967*.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011) "In Search of Attention," *The Journal of Finance*, Vol. 66, No. 5, pp. 1461–1499, October.

——— (2015) "The Sum of All FEARS Investor Sentiment and Asset Prices," *The Review of Financial Studies*, Vol. 28, No. 1, pp. 1–32.

Da, Zhi, Umit G. Gurun, and Mitch Warachka (2014) "Frog in the Pan: Continuous Information and Momentum," *Review of Financial Studies*, Vol. 27, No. 7, pp. 2171–2218, July.

Da, Zhi, Jian Hua, Tim Chih-Ching Hung, and Lin Peng (2025) "Market Returns and a Tale of Two Types of Attention," *Management Science*.

Da, Zhi, Qianqiu Liu, and Ernst Schaumburg (2014) "A closer Look at the Short-Term Return Reversal," *Management Science*, Vol. 60, No. 3, pp. 658–674, March.

Drake, Michael S, Jared Jennings, Darren T Roulstone, and Jacob R Thornock (2017) "The Comovement of Investor Attention," *Management Science*, Vol. 63, No. 9, pp. 2847–2867.

Du, Jun, Dashan Huang, Yu-Jane Liu, Yushui Shi, Avanidhar Subrahmanyam, and Huacheng Zhang (2025) "Nominal Prices, Retail Investor Participation, and Return Momentum," *Management Science*, p. mnsc.2023.01423, July.

Gargano, Antonio and Alberto G Rossi (2018) "Does It Pay to Pay Attention?" *The Review of Financial Studies*, Vol. 31, No. 12, pp. 4595–4649, December.

Gervais, Simon, Ron Kaniel, and Dan H. Mingelgrin (2001) "The High-Volume Return Premium," *The Journal of Finance*, Vol. 56, No. 3, pp. 877–919, June.

Hou, K (2009) "A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum," *Working Paper, Available at SSRN 976394*.

Hu, Grace Xing, Jun Pan, and Jiang Wang (2021) "Chinese Capital Market: An Empirical Overview," *Critical Finance Review*, Vol. 10, No. 2, pp. 125–206.

Huberman, Gur (2001) "Familiarity Breeds Investment," *The Review of Financial Studies*, Vol. 14, No. 3, pp. 659–680.

Jiang, Lei, Jinyu Liu, Lin Peng, and Baolian Wang (2022) "Investor Attention and Asset Pricing Anomalies," *Review of Finance*, Vol. 26, No. 3, pp. 563–593.

Kempf, Elisabeth, Alberto Manconi, and Oliver Spalt (2017) "Distracted shareholders and Corporate Actions," *The Review of Financial Studies*, Vol. 30, No. 5, pp. 1660–1695, May.

Levy, Hagit, Alberto G Rossi, Ron Shalev, and Emanuel Zur (2024) "What's in Investors' Information Set?" *Working Paper, Available at SSRN 4068922*.

Li, Jun and Jianfeng Yu (2012) "Investor Attention, Psychological Anchors, and Stock Return Predictability," *Journal of Financial Economics*, Vol. 104, No. 2, pp. 401–419.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan (2019) "Size and Value in china," *Journal of Financial Economics*, Vol. 134, No. 1, pp. 48–69, October.

Merton, Robert C. (1987) "A simple Model of Capital Market Equilibrium with Incomplete Information," *The Journal of Finance*, Vol. 42, No. 3, pp. 483–510, July.

Peng, Lin (2005) "Learning with Information Capacity Constraints," *Journal of Financial and Quantitative Analysis*, Vol. 40, No. 2, pp. 307–329, June.

Rabin, Matthew and Joel L Schrag (1999) "First Impressions Matter: A Model of Confirmatory Bias," *The Quarterly Journal of Economics*, Vol. 114, No. 1, pp. 37–82.

Sicherman, Nachum, George Loewenstein, Duane J. Seppi, and Stephen P. Utkus (2016) "Financial Attention," *Review of Financial Studies*, Vol. 29, No. 4, pp. 863–897, April.

Staw, Barry M (1976) "Knee-deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action," *Organizational Behavior and Human Performance*, Vol. 16, No. 1, pp. 27–44.

Tan, Lin, Xiaoyan Zhang, and Xinran Zhang (2023) "Retail and Institutional Investor Trading Behaviors: Evidence From China," *Annual Review of Financial Economics*, Vol. 16.

Tetlock, Paul C. (2011) "All The News That's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies*, Vol. 24, No. 5, pp. 1481–1512, May.

Thaler, Richard (1980) "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior & Organization*, Vol. 1, No. 1, pp. 39–60.

Vozlyublennaia, Nadia (2014) "Investor Attention, Index Performance, and Return Predictability," *Journal of Banking & Finance*, Vol. 41, pp. 17–35.

Yuan, Yu (2015) "Market-Wide Attention, Trading, and Stock Returns," *Journal of Financial Economics*, Vol. 116, No. 3, pp. 548–564, June.

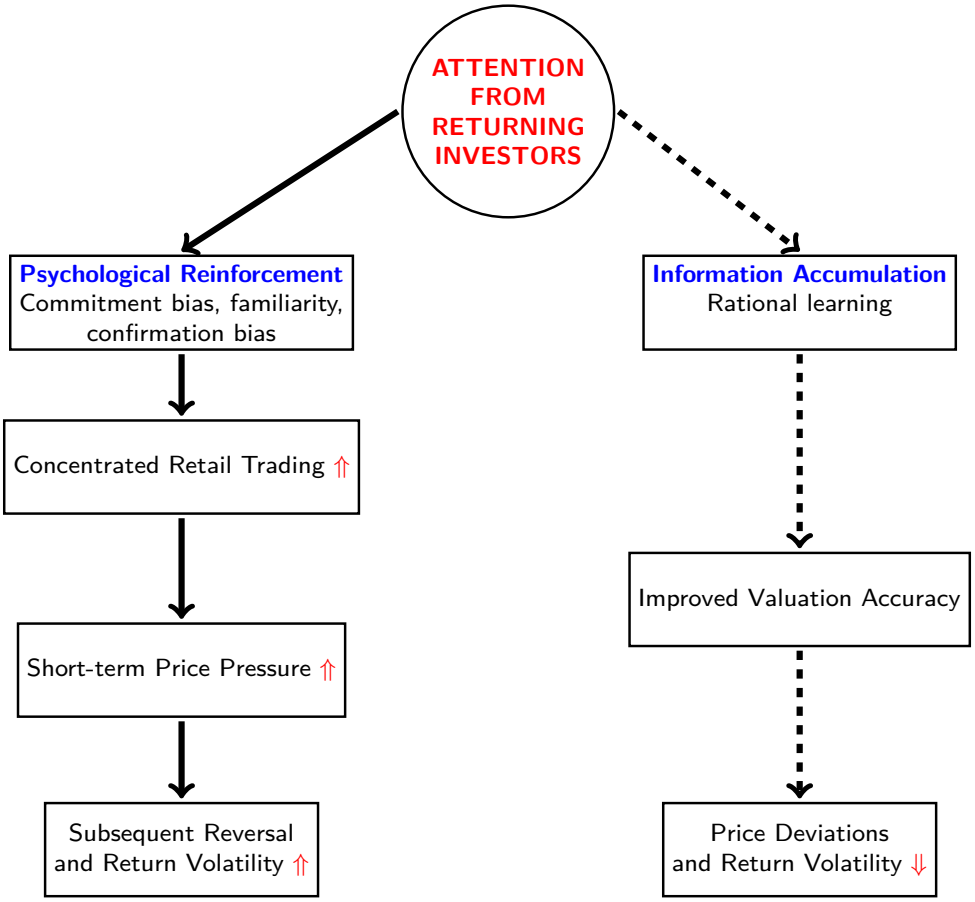


Figure 1: Conceptual Framework: Behavioral Reinforcement versus Rational Learning
This figure illustrates the two interpretations of viewer-base persistence. Under behavioral reinforcement (left path), repeated exposure may make trading intentions more correlated, concentrating temporary retail demand and producing short-term price pressure followed by reversal. Under rational learning (right path), returning viewers may process information more accurately, leading to smaller transient price reactions and weaker subsequent reversal. The two interpretations are not mutually exclusive; the empirical analysis examines which is more consistent with the observed price dynamics (Sections 2–4).

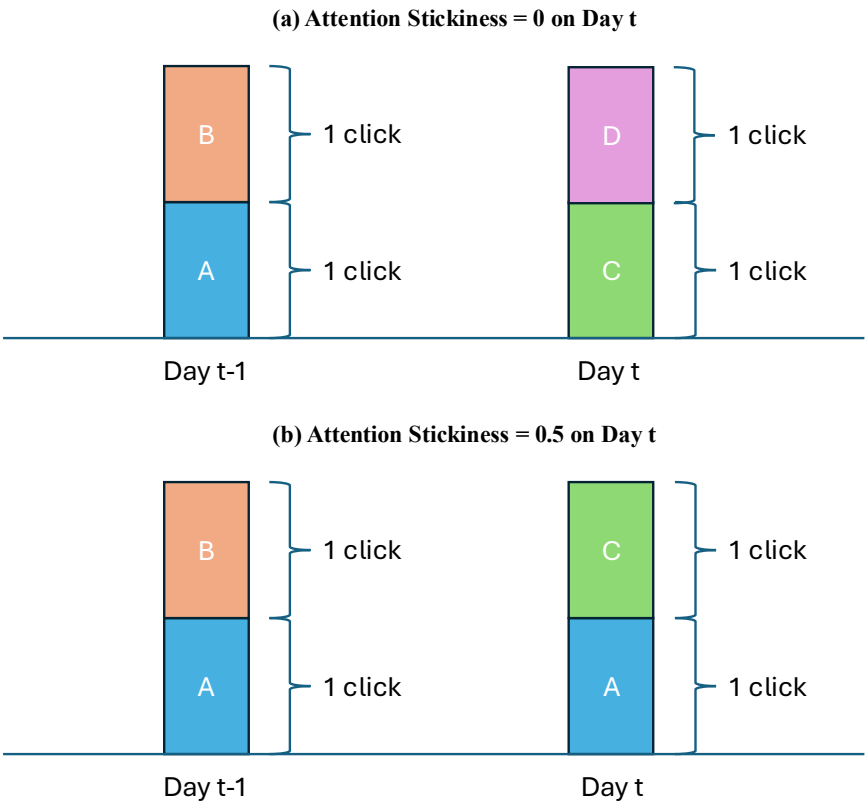


Figure 2: Illustration of the Attention Stickiness Measure
This figure illustrates the attention stickiness measure $STK_{i,t,k}$, defined as the fraction of day-$t-k$ viewers who return to view the same stock page on day t (see Equation 2). Each cell represents one user-day observation; different colors denote different investors. Panel (a) shows a zero-stickiness case: the two users present on day $t-1$ (investors A and B) are fully replaced by new users on day t (investors C and D), so $STK_{i,t,1} = 0/2 = 0$. Panel (b) shows a partial-stickiness case: one of the two prior-day users (investor A) returns alongside one new user (investor C), so $STK_{i,t,1} = 1/2 = 0.5$. Total attention—two visits per day—is identical across panels; only the temporal composition of the viewer set differs.

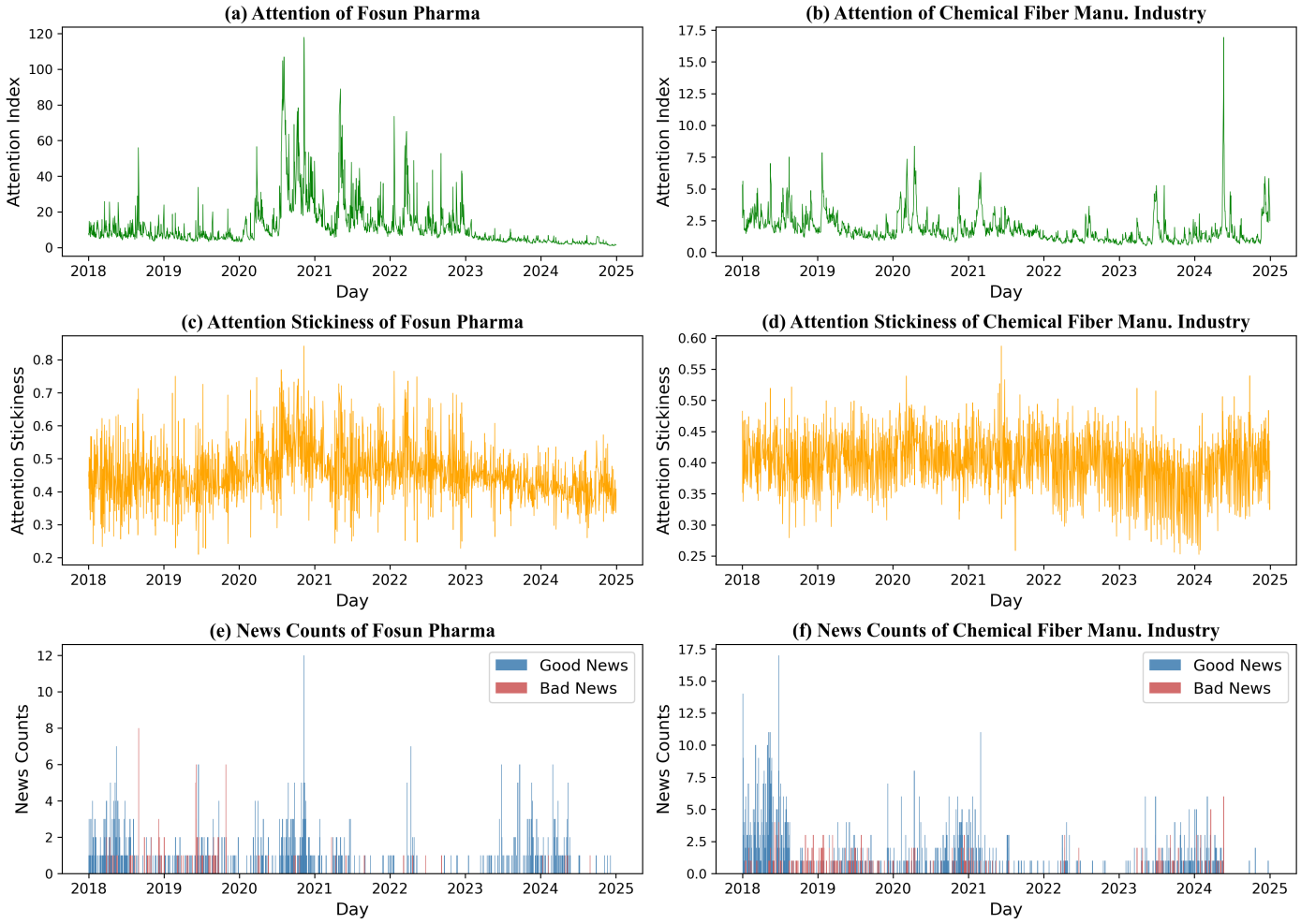


Figure 3: Time-Series Patterns of Attention, Attention Stickiness, and News
This figure plots daily time-series of three measures stacked vertically in each panel: the retail attention index (top row), 1-day attention stickiness ratio (middle row), and daily media article count (bottom row). Left panels report firm-level patterns for Shanghai Fosun Pharmaceutical (stock code 600196); right panels report industry-level patterns for the Chemical Fiber Manufacturing industry. Industry-level attention and attention stickiness are computed via bottom-up equal-weighted aggregation of constituent firms; industry news counts are cross-sectional sums of firm-level article counts. The figure illustrates that attention stickiness (middle row) is substantially smoother and less reactive to news bursts than the aggregate attention index (top row), providing visual evidence that the two series capture distinct dimensions of investor behavior independent of information arrival. The sample period is from January 2018 to December 2024.

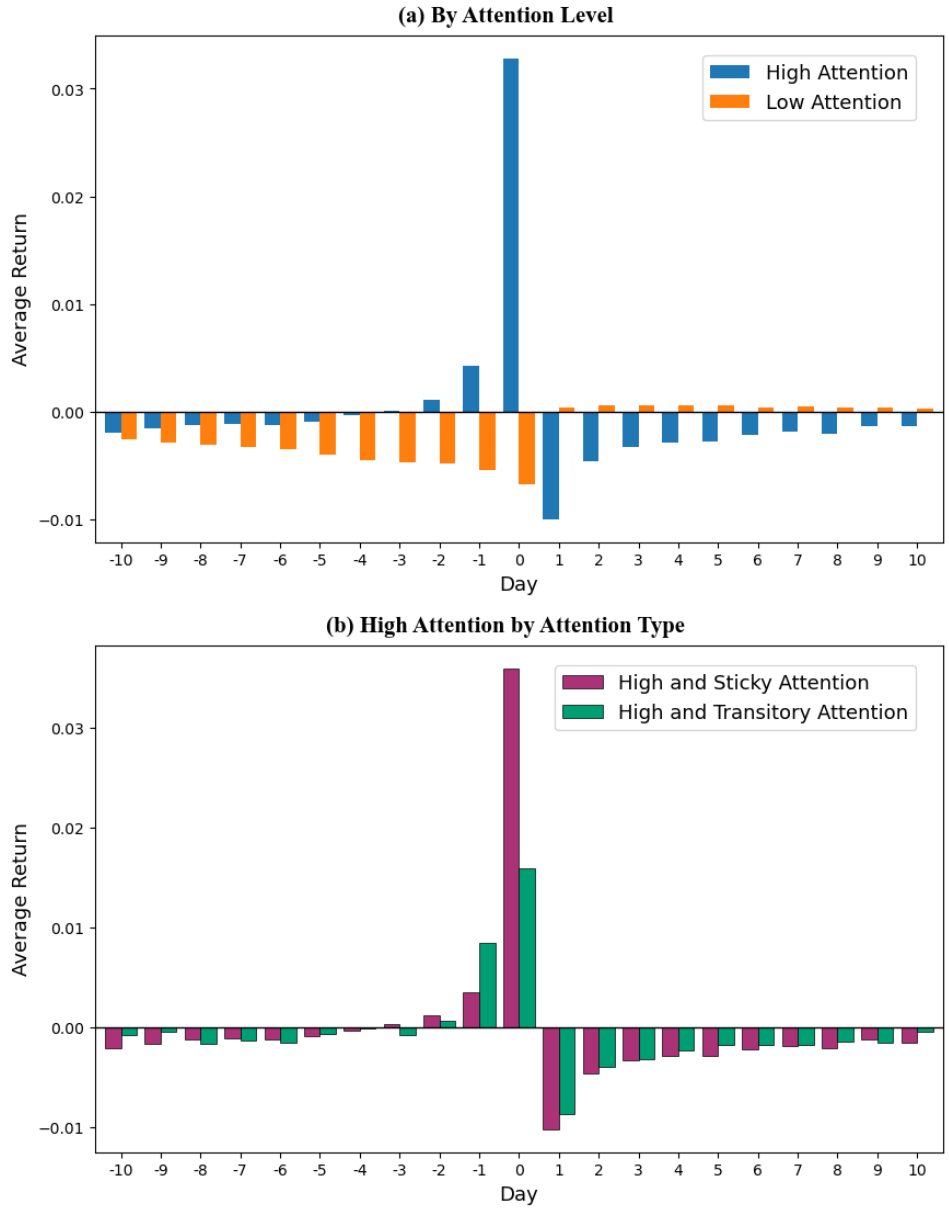


Figure 4: Return Dynamics During High Attention Episodes
This figure reports average stock return dynamics around high abnormal-attention episodes. A high-attention episode is defined as a stock-day on which the stock's abnormal attention ($ATT_{i,t}$, the standardized log-ratio of the daily attention index to its 30-day rolling median) enters the top decile of the full-sample distribution. For each episode, we compute the average holding-period return relative to day 0 over a $[-10, +10]$-day event window. Panel (a) compares the event-window returns of focal high-attention stocks with those of stocks in the bottom abnormal-attention decile on the same day (the benchmark). Panel (b) partitions high-attention episodes into sticky-attention episodes (1-day attention stickiness above the sample median) and transitory-attention episodes (below the median), and compares their return dynamics. The sample period is from January 2018 to December 2024.

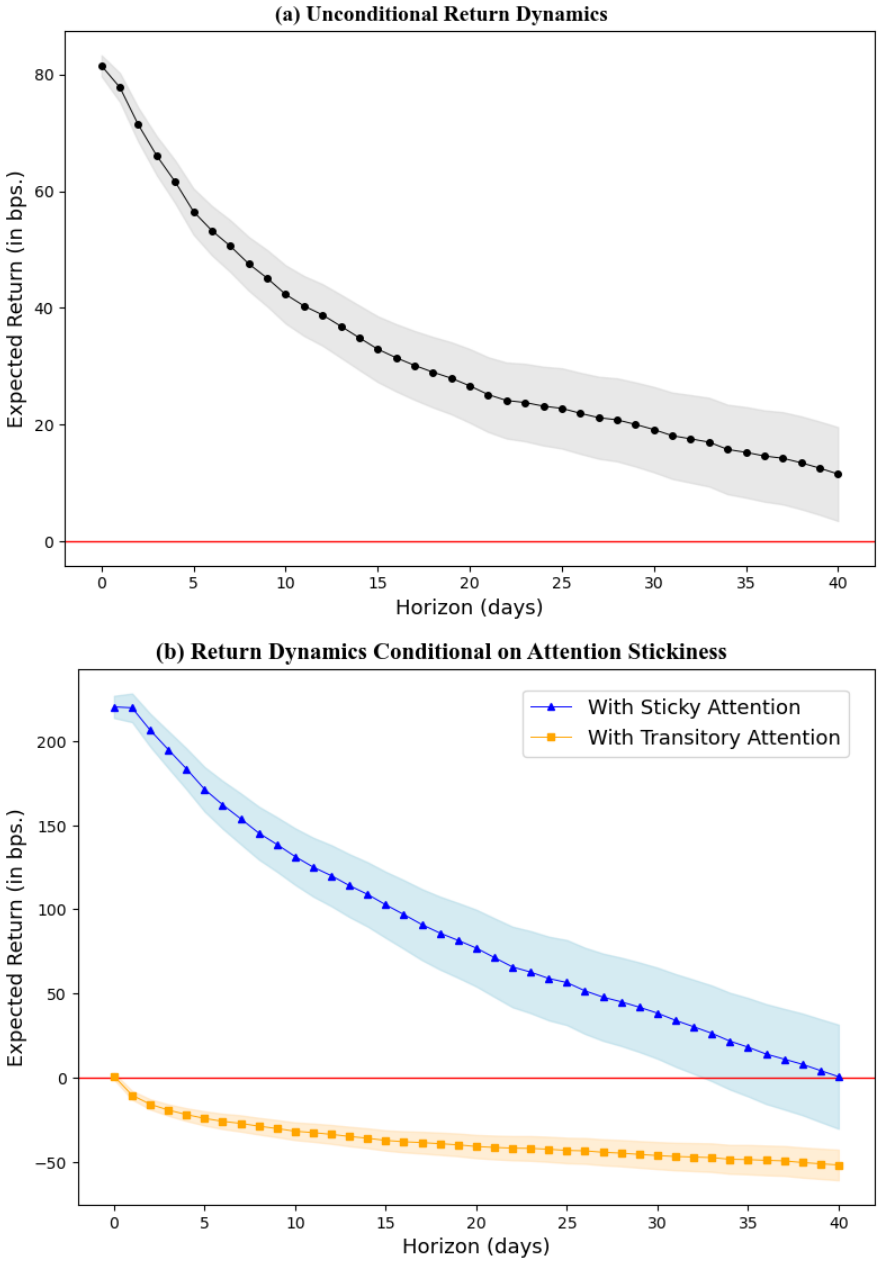


Figure 5: Return Dynamics with Abnormal Attention and Attention Stickiness
This figure reports estimates from panel regressions of return dynamics: $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the cumulative return from the closing price on day $t-1$ to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention index, and $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if stock i's attention stickiness exceeds the panel median on day t, and zero otherwise. We let h range from 0 to 40. Panel (a) plots the unconditional estimates of $\beta_{1,h} + \gamma_h$, representing a one standard deviation shock to abnormal attention, obtained by setting $\beta_{2,h} = \beta_{3,h} = 0$ in the full specification. Panel (b) plots the estimates conditional on attention stickiness, where the square-dotted line shows estimates of $\beta_{1,h} + \gamma_h$, capturing price dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows estimates of $\beta_{1,h} + \beta_{2,h} + \beta_{3,h} + \gamma_h$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. All regressions include firm and time fixed effects. The sample period is from January 2018 to December 2024.

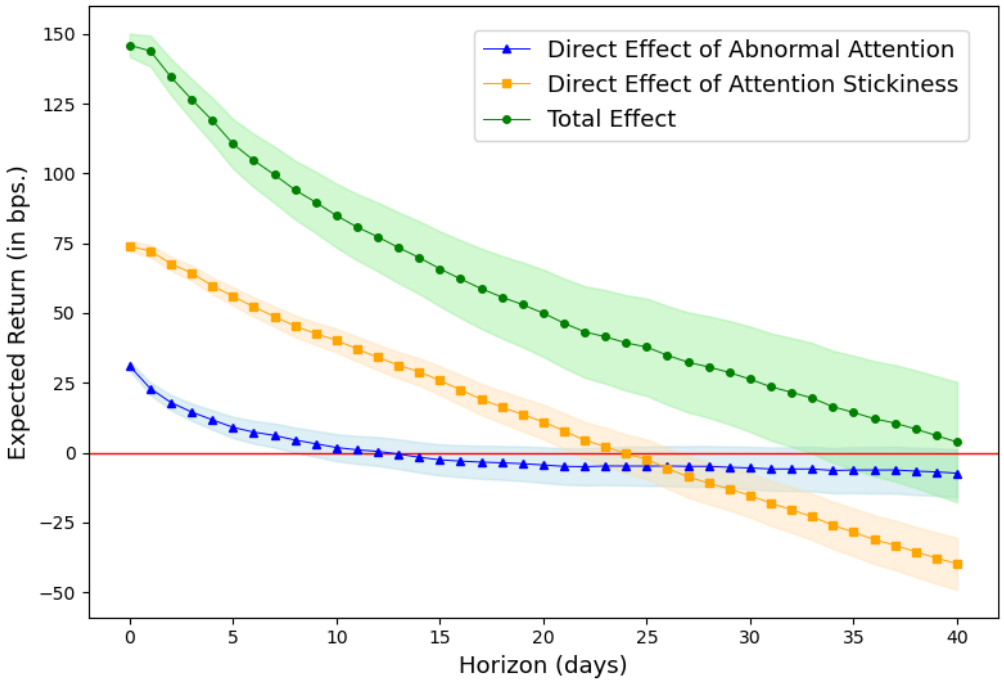


Figure 6: Price Dynamics with Continuous Attention Stickiness
This figure reports estimates from panel regressions of return dynamics: $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$, where the dependent variable is the holding period return from the closing price on time $t-1$ to that on $t+h$. Independent variables include normalized abnormal attention, normalized 1-day attention stickiness, and their interaction. One-day lagged return is added as a control variable. The blue triangle-dotted line shows the estimates of $\beta_{1,h}$, representing the marginal effect of a standard deviation shock to abnormal attention when stickiness is at unconditional mean. The orange square-dotted line shows the estimates of $\beta_{2,h}$, indicating the marginal effect to attention stickiness when abnormal attention is at unconditional mean. The green round-dotted line shows the estimates of $\beta_{1,h} + \beta_{2,h} + \beta_{3,h}$, representing the cumulative marginal effect to both abnormal attention and attention stickiness.

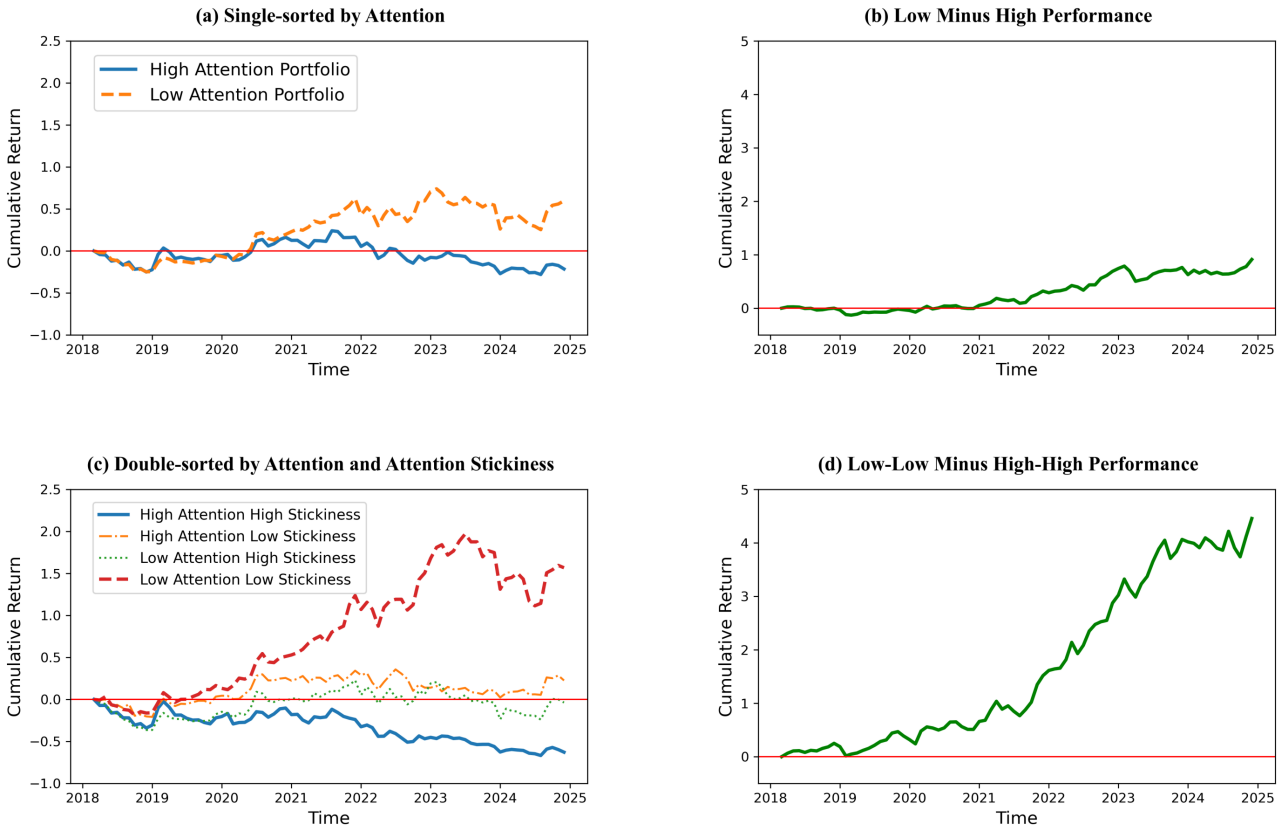


Figure 7: Cumulative Returns of Attention-Sorted Portfolios
This figure shows cumulative returns for value-weighted portfolios formed on abnormal retail attention and attention stickiness. Stocks are sorted into quintiles using the prior month's average $ATT_{i,t}$ (standardized log-abnormal retail attention) or 1-day attention stickiness ratio, and portfolios are rebalanced monthly. Panel (a) plots cumulative returns for the top and bottom abnormal-attention quintile portfolios. Panel (b) plots the corresponding low-minus-high abnormal-attention long-short portfolio. Panel (c) shows cumulative returns for the corner portfolios of a 5×5 double-sort on prior-month average $ATT_{i,t}$ and attention stickiness. Panel (d) plots the Low-Low minus High-High long-short strategy (performance statistics are reported in Table 6). The sample period is from January 2018 to December 2024.

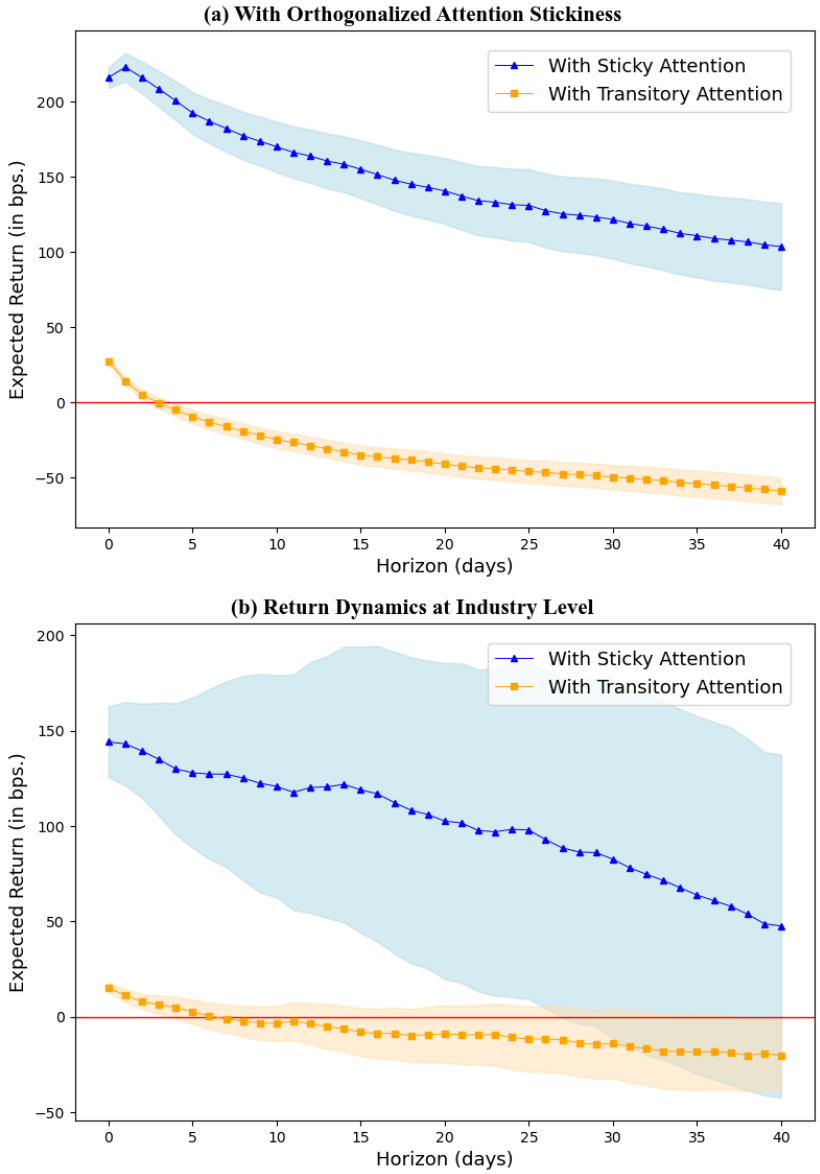


Figure 8: Return Dynamics: Orthogonalized Stickiness and Industry-Level Estimates
This figure reports estimates from panel regressions of return dynamics: $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the cumulative return from the closing price on day $t-1$ to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention index, and $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if the relevant attention stickiness measure exceeds the panel median on day t, and zero otherwise. We let h range from 0 to 40. The square-dotted line shows estimates of $\beta_{1,h}+\gamma_h$, capturing price dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows estimates of $\beta_{1,h}+\beta_{2,h}+\beta_{3,h}+\gamma_h$, representing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. Panel (a) uses the orthogonalized stickiness measure $STK_{i,t}^{\perp}$, constructed by regressing $STK_{i,t}$ on $ATT_{i,t}$, lagged $ATT_{i,t}$, and other firm-specific characteristics, then retaining the intercept plus residual. Panel (b) reports industry-level estimates in which firm-level attention, stickiness, and returns are aggregated to industries via bottom-up equally-weighting. The sample period is from January 2018 to December 2024.

Table 1: Summary Statistics

This table presents summary statistics for the main variables. The sample consists of 3,144 A-share stocks traded on the Shanghai (SSE) and Shenzhen (SZSE) stock exchanges at daily frequency from January 2018 to December 2024. Stocks designated as *Special Treatment* ("ST" or "ST*") or with negative shareholders' equity are excluded. Stocks listed on the ChiNext (Shenzhen) or STAR (Shanghai) boards are also excluded, as they operate under distinct regulatory regimes and firm characteristics. Observations within the initial five trading days of IPO are dropped from the sample. Attention Index, Searching Index, and Viewing Index are each stock's daily share of total platform-wide activity (%). Abnormal Attention ($ATT_{i,t}$) is the standardized log-ratio of the daily attention index to its 30-day rolling median. The k-Day Stickiness Ratio ($STK_{i,t,k}$) is the fraction of users who viewed stock i on day $t-k$ and returned on day t, for $k \in \{1, 7, 30\}$. Market Cap and Total Asset are measured in natural logarithms of CNY.

	Obs.	Mean	SD	Min	Max	10% Pct.	90% Pct.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Attention Indexes							
Attention Index (%)	4,671,745	0.026	0.060	0.000	5.022	0.003	0.055
Searching Index (%)	4,671,745	0.025	0.078	0.000	7.813	0.001	0.053
Viewing Index (%)	4,671,745	0.026	0.060	0.000	5.029	0.003	0.055
Abnormal Attention	4,671,745	0.079	0.655	-4.064	7.537	-0.520	0.851
1-Day Stickiness Ratio	4,671,745	0.386	0.100	0.030	0.866	0.258	0.505
7-Day Stickiness Ratio	4,671,725	0.331	0.106	0.012	0.863	0.199	0.458
30-Day Stickiness Ratio	4,671,745	0.233	0.099	0.006	0.819	0.115	0.358
Panel B: Stock Characteristics							
Leverage	4,486,618	0.391	0.219	0.000	0.998	0.108	0.694
Market Cap	4,671,745	15.868	1.121	12.722	21.907	14.650	17.365
ROA	4,487,250	0.028	0.063	-3.164	3.223	-0.007	0.081
Total Asset	4,487,250	22.546	1.514	17.458	31.468	20.943	24.430
Turnover	4,671,745	0.017	0.024	0.000	0.603	0.003	0.038
Stock Price	4,671,745	15.830	36.835	0.360	2601.000	3.740	29.650
Stock Return	4,671,745	0.000	0.028	-0.105	0.107	-0.029	0.030

Table 2: Drivers of Attention and Attention Stickiness
This table reports OLS estimates from cross-sectional regressions of normalized abnormal attention and attention stickiness on firm-level characteristics. The specification is $Y_{i,t} = \alpha + \beta \times X_{i,t} + \epsilon_{i,t}$. The dependent variable is normalized abnormal attention in Panel A and normalized attention stickiness in Panel B. Newey-West standard errors with the optimal bandwidth are reported. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	(1)	(2)	(3)	(4)	(5)
Panel A: Drivers of Attention					
Stock Return Positive	24.422***	24.355***	24.120***	24.470***	27.356***
	(0.013)	(0.013)	(0.014)	(0.015)	(0.015)
Stock Return Negative	-5.251***	-5.224***	-5.046***	-5.394***	-4.682***
	(0.015)	(0.015)	(0.016)	(0.016)	(0.018)
Good News Dummy		0.095***	0.097***	0.072***	0.073***
		(0.001)	(0.001)	(0.001)	(0.001)
Bad News Dummy		0.104***	0.105***	0.088***	0.095***
		(0.002)	(0.002)	(0.002)	(0.002)
Neutral News Dummy		0.052***	0.054***	0.033***	0.047***
		(0.001)	(0.001)	(0.001)	(0.001)
Turnover			0.477***	0.892***	1.416***
			(0.011)	(0.012)	(0.013)
Leverage				-0.052***	-0.078***
				(0.001)	(0.003)
ROA				-0.009**	-0.036***
				(0.004)	(0.004)
Total Asset (log)				0.023***	0.034***
				(0.000)	(0.001)
Lag Attention	0.710***	0.708***	0.705***	0.698***	0.692***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Adj. R^2	0.768	0.769	0.769	0.768	0.789
Panel B: Drivers of Attention Stickiness					
Stock Return Positive	28.238***	28.211***	26.062***	26.348***	29.021***
	(0.024)	(0.024)	(0.026)	(0.027)	(0.026)
Stock Return Negative	-3.168***	-3.149***	-1.429***	-1.812***	-0.781***
	(0.027)	(0.027)	(0.028)	(0.029)	(0.032)
Good News Dummy		0.041***	0.048***	-0.001	-0.038***
		(0.002)	(0.002)	(0.002)	(0.002)
Bad News Dummy		0.060***	0.064***	0.033***	-0.004
		(0.004)	(0.004)	(0.004)	(0.003)
Neutral News Dummy		0.000	0.008***	-0.037***	-0.077***
		(0.003)	(0.003)	(0.003)	(0.003)
Turnover			4.049***	5.037***	6.850***
			(0.021)	(0.023)	(0.024)
Leverage				-0.137***	-0.294***
				(0.003)	(0.005)
ROA				0.351***	0.385***
				(0.008)	(0.009)
Total Asset (log)				0.041***	0.158***
				(0.000)	(0.002)
Lag STK	-0.019***	-0.020***	-0.041***	-0.051***	-0.064***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Adj. R^2	0.250	0.250	0.257	0.261	0.319
Observations	4,668,606	4,668,606	4,668,606	4,483,487	4,483,487
FE	NO	NO	NO	NO	Firm/Time

Table 3: Return Predictability from Abnormal Retail Attention

This table reports coefficient estimates from panel regressions of future stock returns on abnormal retail attention. The specification is $r_{i,t+1\to t+h} = \beta_h \times ATT_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+h$; Column (1) reports the contemporaneous association, where the dependent variable is the close-to-close return on day t. Panel A uses $ATT_{i,t}$, the standardized log-abnormal retail attention index, as predictor. Panel B uses a high-attention indicator that equals one when stock i's abnormal attention exceeds the median of panel distribution on day t. All regressions include firm and time fixed effects; standard errors are clustered by firm and time. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Same-day	Future Horizons (days)				
	$h=0$	$h=1$	$h=5$	$h=10$	$h=20$	$h=40$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Standardized Abnormal Attention Index as Predictor						
ATT	81.564***	-4.627***	-25.758***	-39.812***	-55.131***	-69.909***
	(0.911)	(0.698)	(1.626)	(2.198)	(2.927)	(3.863)
$r_{i,t-1}$	-0.036***	0.004	-0.014	-0.025**	-0.050***	-0.073***
	(0.005)	(0.004)	(0.009)	(0.011)	(0.015)	(0.016)
Adj. R^2	0.326	0.248	0.257	0.247	0.247	0.256
Panel B: High Abnormal Attention Indicator as Predictor						
$\mathbb{I}^{ATT}$	93.434***	-6.399***	-25.325***	-36.953***	-55.860***	-79.944***
	(1.260)	(0.973)	(2.222)	(3.096)	(4.256)	(5.865)
$r_{i,t-1}$	0.019***	0.001	-0.033***	-0.055***	-0.089***	-0.119***
	(0.005)	(0.004)	(0.009)	(0.011)	(0.015)	(0.016)
Adj. R^2	0.275	0.248	0.256	0.246	0.246	0.255
Observations	4,671,745	4,668,620	4,656,121	4,640,501	4,609,177	4,546,441
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time

Table 4: Return Predictability Conditional on Attention Stickiness

This table reports coefficient estimates from panel regressions of future stock returns on abnormal retail attention, the attention stickiness indicator, and their interaction. The specification is $r_{i,t+1\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+h$; Column (1) reports the contemporaneous association, where the dependent variable is the close-to-close return on day t. Independent variables include normalized abnormal attention, attention stickiness, and their interaction. One-day lagged stock return is added as a control variable. Panel A uses normalized attention stickiness as the predictor, while Panel B replaces the predictor by an indicator, which equal to one if stock i's attention stickiness exceeds the panel median on day t, and zero otherwise. All regressions include firm and time fixed effects; standard errors are clustered by firm and time. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Same-day	Future Horizons (days)				
	$h=0$	$h=1$	$h=5$	$h=10$	$h=20$	$h=40$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Prediction by Normalized Attention Stickiness						
ATT	31.027***	-8.515***	-22.587***	-30.067***	-36.394***	-39.400***
	(0.807)	(0.690)	(1.651)	(2.215)	(3.034)	(4.200)
STK	73.934***	-1.916***	-18.205***	-33.837***	-62.675***	-112.615***
	(0.905)	(0.551)	(1.367)	(1.947)	(3.026)	(4.673)
$ATT \times STK$	40.899***	7.169***	4.618***	2.506**	3.573**	11.451***
	(0.412)	(0.304)	(0.782)	(1.082)	(1.433)	(1.756)
$r_{i,t-1}$	0.071***	0.006*	-0.031***	-0.060***	-0.115***	-0.188***
	(0.004)	(0.004)	(0.010)	(0.011)	(0.015)	(0.016)
Adj. R^2	0.453	0.249	0.257	0.248	0.248	0.258
Panel B: Prediction by Attention Stickiness Indicator						
ATT	0.914	-11.410***	-25.427***	-33.477***	-42.796***	-53.933***
	(0.963)	(0.734)	(1.828)	(2.407)	(3.310)	(4.501)
$\mathbb{I}^{STK}$	107.584***	-1.803**	-28.084***	-53.295***	-93.260***	-160.441***
	(1.283)	(0.785)	(1.943)	(2.769)	(4.299)	(6.629)
$ATT \times \mathbb{I}^{STK}$	111.655***	10.901***	3.391*	-2.510	-6.357*	-2.727
	(1.101)	(0.704)	(1.740)	(2.376)	(3.248)	(4.083)
$r_{i,t-1}$	0.022***	0.004	-0.024**	-0.046***	-0.087***	-0.135***
	(0.004)	(0.004)	(0.009)	(0.011)	(0.015)	(0.016)
Adj. R^2	0.387	0.248	0.257	0.247	0.248	0.257
Observations	4,671,745	4,668,620	4,656,121	4,640,501	4,609,177	4,546,441
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time

Table 5: Characteristics of Portfolios
This table reports average characteristics for value-weighted 5×5 portfolios double-sorted on abnormal retail attention (ATT) and 1-day attention stickiness (STK). Stocks are sorted into quintiles using the prior month's average $ATT_{i,t}$ and $STK_{i,t,1}$; portfolios are held for one month and rebalanced monthly. Panel A reports the average monthly return (%) in the holding month. Panel B reports the average monthly turnover (trading volume divided by shares outstanding) in the holding month. The sample period is from January 2018 to December 2024.

Panel A: Ex-post Average Monthly Return (%)

	Attention Stickiness				
	Q1	Q2	Q3	Q4	Q5
Abnormal Attention					
Q1	1.311	0.977	0.876	0.607	0.206
Q2	1.162	0.654	0.703	0.670	-0.044
Q3	1.057	1.060	0.786	0.664	0.227
Q4	0.410	0.616	0.341	0.042	-0.204
Q5	0.441	0.551	0.252	-0.240	-0.922

Panel B: Ex-post Average Monthly Turnover

	Attention Stickiness				
	Q1	Q2	Q3	Q4	Q5
Abnormal Attention					
Q1	0.133	0.152	0.201	0.248	0.366
Q2	0.098	0.113	0.135	0.165	0.252
Q3	0.099	0.112	0.131	0.166	0.239
Q4	0.112	0.129	0.150	0.189	0.297
Q5	0.178	0.217	0.270	0.339	0.545

Table 6: Long-Short Portfolio Performance

This table reports the performance of a value-weighted long-short strategy that goes long the Low-Attention/Low-Stickiness (Low-Low) portfolio and short the High-Attention/High-Stickiness (High-High) portfolio. Stocks are double-sorted into 5×5 portfolios on the prior month's average abnormal attention and 1-day attention stickiness ratio; portfolios are rebalanced monthly. Panel A reports compounded buy-and-hold excess returns and annualized Sharpe ratios over holding periods of 1 to 6 months. Panel B reports factor alphas from daily time-series regressions of the Low-Low minus High-High portfolio excess return on the CAPM market factor (Column 2), the Fama-French three factors (market, SMB, HML; Column 3), and a four-factor model adding the sentiment factor PMO (Column 4). The regression is $r_t^{LMH} = \alpha + \beta^{MKT} \times (r_t^{MKT} - r_t^f) + \beta^{SMB} \times SMB_t + \beta^{HML} \times HML_t + \beta^{PMO} \times PMO_t + \epsilon_t$; standard errors use the Newey-West formula with seven lags. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

Panel A Buy-and-Hold Performance

	Holding Period (Month)					
	1	2	3	4	5	6
Compounded Excess Return (%)	2.235	4.450	6.615	8.930	11.351	13.739
Standard Deviation (%)	5.633	8.516	9.972	10.839	11.208	11.789
Annualized Sharpe Ratio	1.375	1.280	1.327	1.427	1.569	1.648

Panel B Alpha Tests

	Raw Return	CAPM	3-Factor	4-Factor	Exclude Illiquid
α	2.235***	2.302***	2.052***	1.455***	1.466***
	(0.587)	(0.565)	(0.559)	(0.528)	(0.506)
β^{MKT}		-35.190***	-29.101**	-34.011***	-26.370***
		(11.309)	(11.769)	(9.985)	(10.055)
β^{SMB}			27.299	16.108	15.298
			(16.843)	(14.323)	(14.379)
β^{HML}			31.065*	15.033	15.605
			(17.462)	(14.913)	(14.938)
β^{PMO}				87.128***	88.647***
				(15.650)	(15.620)
Observations	82	82	82	82	82
R^2		0.107	0.151	0.390	0.368

Table 7: Heterogeneity: Nominal Price and Institutional Ownership
This table reports coefficient estimates from the return-predictability specification of Table 4, estimated separately for cross-sectional subgroups. The specification is $r_{i,t+1\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized abnormal retail attention, and $STK_{i,t}$ denotes the standardized attention stickiness ratio. Nominal price proxies for retail investor participation (low-priced stocks disproportionately attract retail attention); institutional ownership proxies for the presence of sophisticated arbitrageurs who can correct mispricing. Stocks are sorted into quintiles by prior-month nominal stock price (Panel A) and by prior-year institutional ownership share (Panel B). We report estimates for the highest and lowest quintiles and their difference (High minus Low). All regressions include firm and time fixed effects; standard errors are clustered by firm and calendar day. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Same-day	Future Horizons (days)				
	$h=0$	$h=1$	$h=5$	$h=10$	$h=20$	$h=40$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sort by Nominal Price						
High Nominal Price						
ATT	37.167***	-8.559***	-25.789***	-32.839***	-44.447***	-53.388***
STK	79.042***	-0.836	-13.011***	-25.691***	-45.079***	-79.999***
$ATT \times STK$	38.181***	5.554***	5.160***	5.453***	8.079***	18.348***
Low Nominal Price						
ATT	27.169***	-8.303***	-23.387***	-35.764***	-49.135***	-64.252***
STK	72.941***	-0.735	-17.572***	-33.689***	-62.073***	-115.636***
$ATT \times STK$	41.018***	7.500***	3.607**	-0.441	0.479	7.959**
High minus Low						
ATT	9.998***	-0.256	-2.402	2.925	4.689	10.864
STK	6.101***	-0.101	4.561	7.998*	16.994**	35.637***
$ATT \times STK$	-2.837***	-1.946***	1.553	5.894**	7.600**	10.389**
Panel B: Sort by Institutional Ownership						
High Institutional Ownership						
ATT	32.118***	-6.370***	-16.997***	-24.723***	-29.116***	-31.462***
STK	66.312***	-1.751***	-15.699***	-27.557***	-51.019***	-91.217***
$ATT \times STK$	35.067***	5.963***	5.781***	4.556**	3.410	6.908**
Low Institutional Ownership						
ATT	27.544***	-11.786***	-31.801***	-43.858***	-53.963***	-54.176***
STK	79.094***	-2.174***	-20.102***	-36.578***	-67.393***	-118.157***
$ATT \times STK$	43.496***	8.357***	2.910	-1.842	-1.238	5.728
High minus Low						
ATT	4.574***	5.416***	14.805***	19.135***	24.847***	22.714**
STK	-12.782***	0.423	4.404*	9.022**	16.374**	26.940**
$ATT \times STK$	-8.429***	-2.394***	2.871	6.397**	4.648	1.180

Online Appendix to

"Sticky Attention and Stock Returns"

(not for publication)

This Online Appendix presents additional results, robustness checks, and supplementary figures and tables that complement the analysis in the main paper.

Contents

A Additional Results **A-3**

A.1 Variable Definitions and Correlations A-3

A.2 Orthogonalization of Attention Stickiness A-3

A.3 Event Study with Searching and Viewing Sub-Indexes A-3

B Additional Robustness Checks **A-5**

B.1 Quarterly Report Disclosures A-5

B.2 Alternative Attention Proxies A-5

B.3 Alternative Stickiness Horizons A-6

B.4 Alternative Classification Cutoffs A-6

B.5 Cross-Sectional Heterogeneity A-7

B.6 Functional Form A-7

B.7 Alternative Abnormal-Attention Windows A-8

C Impact of Attention Stickiness on Stock Volatility **A-9**

C.1 Unconditional Volatility Response A-9

C.2 Decomposition by Attention Stickiness A-10

A Additional Results

This appendix collects supporting material and a supplementary result referenced in the main text.

A.1 Variable Definitions and Correlations

Table A.1 defines all variables used in the analysis, grouping the attention and stickiness measures constructed from the RoyalFlush platform (Panel A) and the stock and firm fundamentals (Panel B). Table A.2 reports pairwise correlations among the key variables. Two patterns support the interpretation of stickiness as a distinct dimension of attention. First, abnormal attention and the stickiness ratio are only modestly correlated (0.25), and the orthogonalized stickiness measure is nearly uncorrelated with attention (0.09), so stickiness is not a relabeling of attention intensity. Second, stickiness correlates positively with the contemporaneous return (0.38) but negatively with the lagged return (-0.14), consistent with the evidence in Table 2 that stickiness reflects same-day demand rather than return chasing.

TABLES A.1 AND A.2 ABOUT HERE

A.2 Orthogonalization of Attention Stickiness

Table A.3 reports the first-stage regression used to construct the orthogonalized stickiness measure $STK^{\perp}_{i,t}$, which regresses each k-day stickiness ratio on attention, its lag, lagged return, log market capitalization, turnover, and news indicators. Observable characteristics explain only a limited share of stickiness variation (adjusted R^2 of 0.28 for the 1-day ratio and less at longer horizons), so the orthogonalized measure retains most of the original variation. This independence underlies the stability of the return results to orthogonalization documented in Section 4.3.

TABLE A.3 ABOUT HERE

A.3 Event Study with Searching and Viewing Sub-Indexes

Section 4.1 shows that high-attention episodes, identified from the composite attention index, are followed by short-term overpricing and a subsequent reversal. Figure A.1 confirms that this baseline

pattern does not depend on how the composite index combines its two underlying signals. We repeat the event study separately for the searching sub-index (Panel a) and the viewing sub-index (Panel b), classifying a high-attention episode as a stock-day in the top decile of the relevant abnormal measure and comparing focal stocks with bottom-decile stocks on the same day. Both sub-indexes reproduce the baseline dynamics: focal stocks earn a sharp positive return around the event day, which reverses over the following days. Because the overreaction appears for each raw activity signal on its own, the attention-driven price pattern documented in the main text reflects shifts in retail monitoring rather than the particular weighting embedded in the composite index.

FIGURE A.1 ABOUT HERE

B Additional Robustness Checks

The sticky-versus-transitory contrast holds across seven robustness checks, each targeting a distinct identification concern: (i) event contamination by quarterly disclosure windows, (ii) sensitivity to the choice of attention proxy, (iii) sensitivity to the horizon used to define stickiness, (iv) sensitivity to the stickiness cutoff, (v) cross-sectional heterogeneity in where the effects concentrate, (vi) sensitivity to the functional form of both attention and stickiness, and (vii) sensitivity to the rolling window used to compute abnormal attention.

B.1 Quarterly Report Disclosures

We examine whether the sticky-attention effects are driven by firms' quarterly report disclosures. Similar to U.S. earnings announcements, quarterly reports are salient firm-specific events in China because listed firms must publicly disclose recent performance.[15] Online Appendix Figure A.2 compares expected return dynamics on disclosure days (Panel (a)) versus other trading days (Panel (b)). The expected return paths for sticky and transitory attention are qualitatively similar across the two subsets: episodes classified as sticky attention exhibit markedly stronger contemporaneous price pressure and a more persistent expected return gap than transitory attention episodes. These results indicate that our baseline stickiness effects are not mechanically driven by quarterly report disclosures.

FIGURE A.2 ABOUT HERE

B.2 Alternative Attention Proxies

We assess robustness to alternative proxies for retail attention by re-estimating specification 4 with *searching* and *viewing* measures as predictors. Online Appendix Figure A.3 reports expected return dynamics: Panel (a) for searching intensity and Panel (b) for page-view intensity. Across both proxies, the baseline patterns are qualitatively unchanged. Stocks with sticky attention display stronger contemporaneous expected returns and a persistent gap relative to transitory attention ones. Columns (2) and (3) of Online Appendix Table A.4 report the corresponding ten-day-ahead coefficient estimates. Our conclusions do not hinge on the specific attention metric used.

[15] Filing deadlines are: April 30 for Q1 reports, August 31 for Q2 (Half-Year) reports, October 31 for Q3 reports, and April 30 of the following year for Q4 (Annual) reports.

FIGURE A.3 ABOUT HERE

B.3 Alternative Stickiness Horizons

We verify robustness to longer-horizon definitions of attention stickiness. Online Appendix Figure A.4 re-estimates the specification using 7-day and 30-day stickiness ratios as identifiers.[16] The results remain qualitatively unchanged, with similar magnitudes and significance levels, indicating limited sensitivity to the horizon choice; columns (4) and (5) of Online Appendix Table A.4 report the corresponding ten-day-ahead coefficient estimates. A related measurement concern is that the k-day ratio ignores revisits within the window (e.g., activity on day $t-5$), potentially introducing attenuation bias. The summary statistics in Table 1 show that the 1- and 7-day averages are close, and the 30-day average is lower but still substantial, suggesting that within-window revisits frequently co-occur with revisits at day $t-k$, limiting bias. The sizable gaps between sticky and transitory attention states under all three horizon definitions reinforce this conclusion.

FIGURE A.4 ABOUT HERE

B.4 Alternative Classification Cutoffs

We vary the cutoff used to classify sticky versus transitory attention. We re-define states using cross-sectional extremes: top/bottom 20% and, more stringently, top/bottom 10% of attention stickiness within each day. Stocks above the upper cutoff are labeled with sticky attention, and those below the lower cutoff are categorized as transitory. As shown in Online Appendix Figure A.5, the expected return gap between the two groups widens under these tighter cutoffs: sticky-attention episodes display larger contemporaneous expected returns and a wider persistent gap, whereas transitory-attention episodes remain muted.

FIGURE A.5 ABOUT HERE

[16] For $k \in \{7, 30\}$, the k-day attention stickiness ratio is the fraction of users who viewed the same stock page on both days t and $t-k$.

B.5 Cross-Sectional Heterogeneity

We examine cross-sectional heterogeneity to gauge where the stickiness-driven dynamics are most pronounced. We sort stocks into quintiles by fundamentals, trading characteristics, and China-specific market conditions (short-sale eligibility and dual-listing in Hong Kong) and rebalance monthly. Online Appendix Table A.5 reports return predictability at the ten-day horizon for the top versus bottom quintile of each characteristic. Panel A shows stronger stickiness effects among smaller and less liquid stocks; market-to-book ratio generates no significant difference. Panel B indicates little variation across past return or volatility sorts, with larger effects for high-turnover stocks. Panel C shows that the stickiness effect is larger where short-sale constraints bind and for stocks without a Hong Kong dual-listing, consistent with the limits-of-arbitrage interpretation: where arbitrageurs face barriers, mispricing driven by sticky attention persists longer and generates larger reversals.

TABLE A.5 ABOUT HERE

B.6 Functional Form

We verify that the sticky–transitory asymmetry does not depend on the functional form of the attention measure. Our baseline uses continuous $ATT_{i,t}$ interacted with stickiness; a natural concern is whether the linear scaling of the continuous index drives the results. Online Appendix Table A.6 replaces both $ATT_{i,t}$ and $STK_{i,t}$ with above-median indicators, yielding a fully non-parametric specification. The interaction $\hat{\beta}_3 = 145.59$ bps at $h = 0$ confirms that the asymmetry is large and significant under this binary-binary design. At longer horizons, the direct stickiness indicator effect ($\hat{\beta}_2 = -56.84$ bps at $h = 10$) and the interaction ($\hat{\beta}_3 = -13.73$ bps at $h = 10$) are both significant and negative, consistent with the continuous-specification and indicator-specification results in Table 4. The stickiness effects are therefore not an artifact of the distributional assumptions embedded in either the continuous attention or stickiness measures.

TABLE A.6 ABOUT HERE

B.7 Alternative Abnormal-Attention Windows

We confirm that the sticky–transitory asymmetry does not depend on the rolling window used to define abnormal attention. The baseline standardizes each stock's attention against its own 30-trading-day median (equation 1); a shorter window mechanically magnifies measured shocks while a longer one smooths them away, so the window length is a researcher degree of freedom that could in principle drive the results. We re-estimate the continuous specification 4, recomputing $ATT_{i,t}$ against 40-, 20-, and 10-trading-day medians in Panels A, B, and C of Online Appendix Table A.7. The stickiness-driven results are stable across all three windows: the same-day stickiness premium ($\hat{\beta}_2$ at $h = 0$) ranges from 66.6 to 75.4 bps, its amplifying interaction with attention ($\hat{\beta}_3$ at $h = 0$) from 38.4 to 41.6 bps, and the forty-day reversal in the stickiness level ($\hat{\beta}_2$ at $h = 40$) from -107.5 to -130.7 bps, each close to the baseline estimates of 73.9, 40.9, and -112.6 bps (Panel A of Table 4). The one coefficient that moves materially is the same-day attention loading itself, which rises from 25.6 to 55.5 bps as the benchmark shortens, a mechanical consequence of a shorter baseline making any given attention level appear more abnormal. The price dynamics we attribute to attention stickiness are therefore a property of investor behavior, not of the benchmark used to scale attention shocks.

TABLE A.7 ABOUT HERE

Across all seven robustness dimensions, the sticky-versus-transitory contrast remains both economically and statistically meaningful. Attention stickiness captures a systematic feature of retail investor behavior that is robust to the choice of attention proxy, the rolling window used to compute abnormal attention, the stickiness horizon, the classification cutoff, sample composition, and functional form.

C Impact of Attention Stickiness on Stock Volatility

Return dynamics capture only one dimension of how attention shapes prices; conditional volatility provides an independent test of the same mechanism. Under rational information-based theories, attention reflects information arrival and volatility rises proportionally to information content, roughly symmetric across sticky and transitory attention (Andrei and Hasler, 2015). Behavioral demand concentration predicts the opposite asymmetry: if returning investors trade in correlated directions, demand concentration amplifies both price pressure and realized uncertainty, while newcomers with heterogeneous beliefs trade in offsetting directions, leaving volatility unchanged or reduced. A volatility asymmetry between the two states would therefore provide additional evidence consistent with demand concentration rather than information processing. We estimate the volatility response to attention shocks, conditioning on attention stickiness, to test these predictions.

Following Andrei and Hasler (2015), we estimate conditional volatility using a GARCH(1,1) model. The dependent variable $\Delta Vol_{i,t+h}$ denotes the change in volatility from day t to $t+h$. We estimate:

$$\Delta Vol_{i,t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}, \quad \text{(A.1)}$$

where $\alpha_{i,h}$ and $\delta_{t,h}$ are firm and time fixed effects. The lagged return control included in the return-dynamics specification (4) is omitted here because the dependent variable is a volatility change rather than a return level. We first report the unconditional attention effect (setting $\beta_{2,h} = \beta_{3,h} = 0$), then the full stickiness decomposition.

C.1 Unconditional Volatility Response

Panel (a) of Figure A.6 plots the expected volatility change following a one-standard-deviation attention shock. The expected change peaks near 9 bps one day after the shock (h=1) and turns negative within approximately four trading days, reaching about -24 bps by day 40: attention spikes inflate conditional uncertainty in the short run, but as the episode fades, volatility falls below its pre-shock baseline. Panel A of Table A.8 reports the underlying coefficients: $+9.23$ bps at $h = 1$, -1.14 bps at $h = 5$, -11.07 bps at $h = 10$, and -23.79 bps at $h = 40$.

FIGURE A.6 AND TABLE A.8 ABOUT HERE

C.2 Decomposition by Attention Stickiness

The conditional decomposition (Panel B of Table A.8 and Panel (b) of Figure A.6) shows that stickiness amplifies the volatility response to attention. A one-standard-deviation attention shock raises next-day conditional volatility by $\hat{\beta}_{1,1} = 3.71$ bps at average stickiness; the positive interaction $\hat{\beta}_{3,1} = 6.57$ bps means the response rises with stickiness, reaching $\hat{\beta}_{1,1} + \hat{\beta}_{3,1} = 10.28$ bps one standard deviation above the mean. A direct volatility premium operates alongside the interaction: $\hat{\beta}_{2,1} = 6.12$ bps, reflecting elevated conditional variance for high-stickiness stocks even holding the attention shock fixed.

Both channels persist across horizons but weaken as the episode fades. The amplification ($\hat{\beta}_{3,h} > 0$) and the stickiness premium ($\hat{\beta}_{2,h} > 0$) remain positive at every horizon, the latter declining from 6.12 bps at $h = 1$ to 1.52 bps at $h = 40$. The attention effect at average stickiness ($\hat{\beta}_{1,h}$) turns negative beyond day 1 (-6.79 bps at $h = 5$, -26.55 bps at $h = 40$), mirroring the unconditional reversal: conditional variance first rises and then falls below its pre-shock baseline.

The volatility and return results admit a common explanation: if returning investors direct correlated order flow, demand concentration amplifies both price pressure and realized volatility. The same feature of the viewer set, whether the same investors return, is associated with the second moment ($\hat{\beta}_2 > 0$, $\hat{\beta}_3 > 0$) just as it is associated with the first-moment overreaction. The joint pattern is consistent with attention stickiness capturing a behavioral dimension of retail activity rather than a residual of aggregate attention intensity.

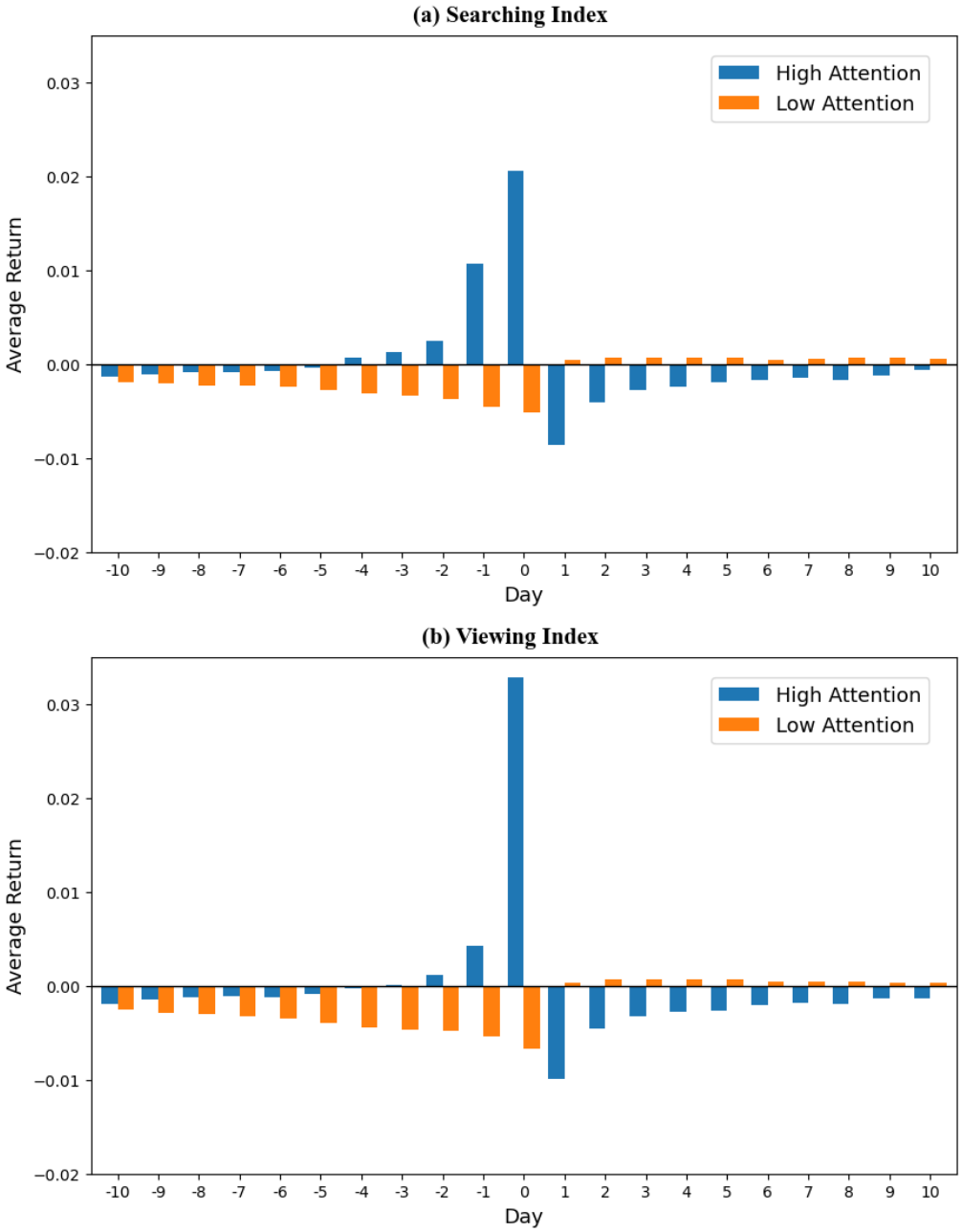


Figure A.1: Return Dynamics During High Attention Episodes: Searching and Viewing Indexes
This figure replicates Figure 4 using the searching and viewing sub-indexes as alternative attention measures. A high-attention episode is defined as a stock-day on which the stock's abnormal searching (Panel a) or viewing (Panel b) index enters the top decile of the full-sample distribution. For each episode, the y-axis plots the average holding-period return relative to day 0 over a $[-10, +10]$-day event window. Panels (a) and (b) compare the event-window returns of focal stocks with those of stocks in the bottom decile of the respective abnormal attention measure on the same day. The sample period is from January 2018 to December 2024.

Table A.4: Return Predictability: Robustness to Alternative Specifications
This table reports coefficient estimates from the return-predictability specification of Table 4 under five alternative settings. The specification is $r_{i,t\to t+10} = \beta_{1,h}\times ATT_{i,t}+\beta_{2,h}\times STK_{i,t}+\beta_{3,h}\times ATT_{i,t}\times STK_{i,t}+\gamma_h\times r_{i,t-1}+\alpha_{i,h}+\delta_{t,h}+\epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+10$. $ATT_{i,t}$ is the standardized log-abnormal retail attention index, and $STK_{i,t}$ denotes the standardized attention stickiness ratio. Column (1) repeats the baseline estimates from Table 4 for reference. Columns (2) and (3) replace the composite attention index with the searching and viewing sub-indexes, respectively. Columns (4) and (5) replace the 1-day stickiness ratio with 7- and 30-day stickiness ratios ($STK_{i,t,7}$ and $STK_{i,t,30}$). All regressions include firm and time fixed effects; standard errors are clustered by firm and calendar day. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

		Alternative Attention Metrics		Alternative Attention Stickiness	
	Baseline Results	Searching	Viewing	7-day	30-day
	(1)	(2)	(3)	(4)	(5)
ATT	-30.067***	-21.284***	-30.064***	-34.300***	-20.698***
	(2.215)	(1.840)	(2.215)	(2.315)	(2.555)
STK	-33.837***	-37.557***	-33.834***	-10.033***	-13.956***
	(1.947)	(1.816)	(1.947)	(1.865)	(2.240)
$ATT \times STK$	2.506**	-1.278	2.507**	-2.248**	-10.771***
	(1.082)	(1.212)	(1.082)	(1.028)	(1.188)
$r_{i,t-1}$	-0.060***	-0.073***	-0.060***	-0.020*	-0.019*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
$AdjR^2$	0.248	0.249	0.248	0.247	0.247
Observations	4,640,501	4,453,961	4,640,501	4,640,481	4,640,501
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time

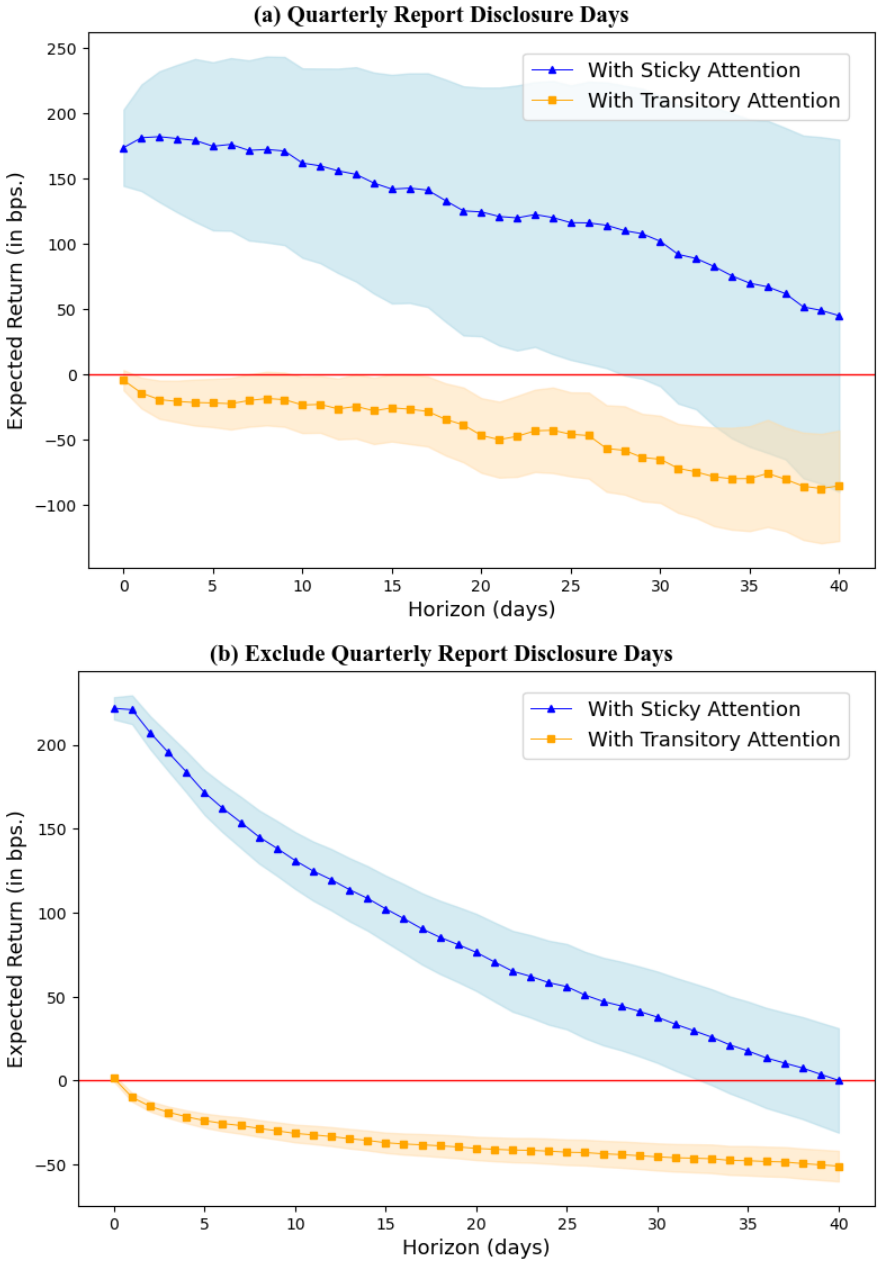


Figure A.2: Return Dynamics Conditional on Quarterly Report Dates
This figure reports estimates from the baseline return dynamics specification (Figure 5) estimated separately on quarterly disclosure dates and non-disclosure dates. The specification is $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the holding period return from the closing price on day $t-1$ to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention, and $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if stock i's attention stickiness exceeds the panel median on day t, and zero otherwise. We let h range from 0 to 40. The square-dotted line shows $\beta_{1,h} + \gamma_h$, capturing return dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows $\beta_{1,h} + \beta_{2,h} + \beta_{3,h} + \gamma_h$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. Panel (a) restricts to quarterly earnings and report disclosure dates; Panel (b) excludes those dates. The sample period is from January 2018 to December 2024.

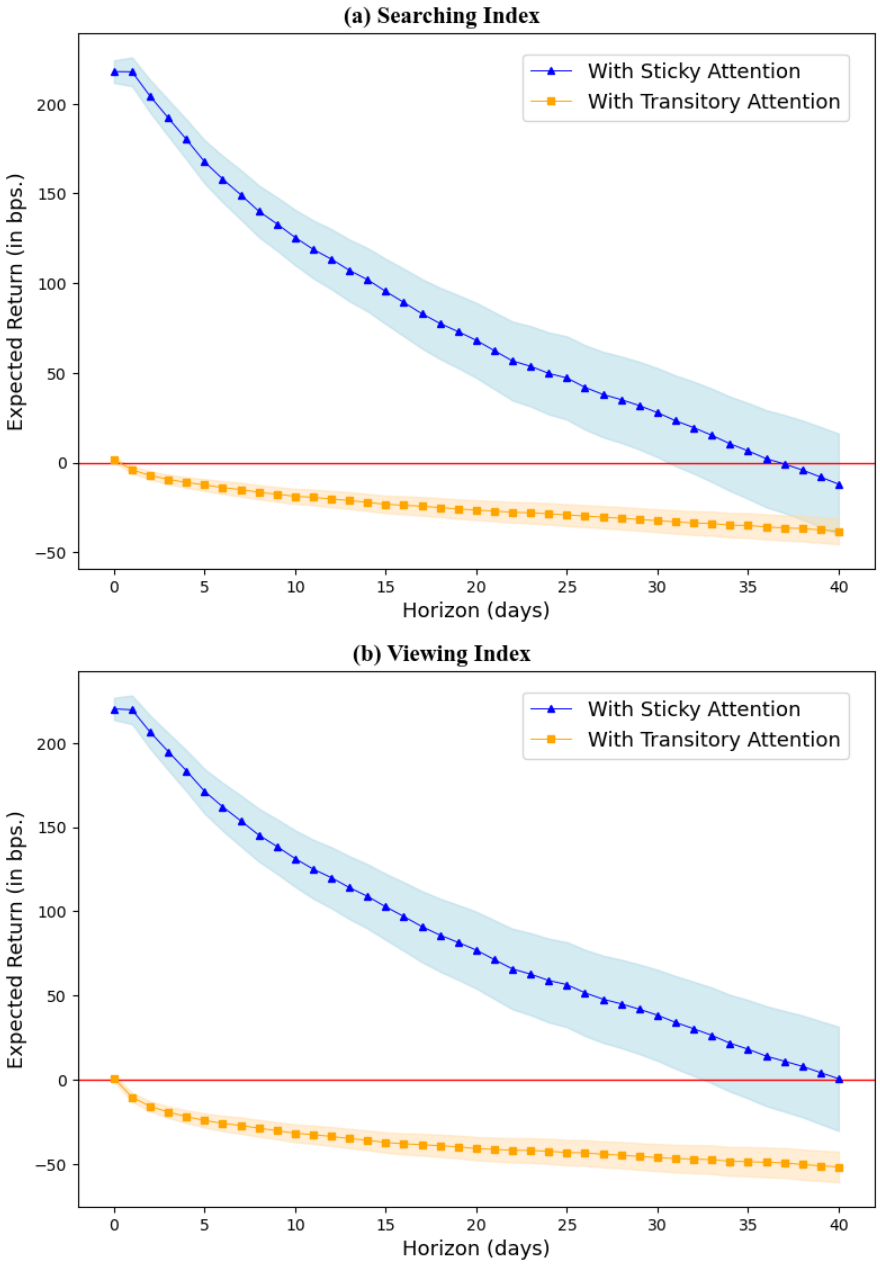


Figure A.3: Return Dynamics Conditional on Attention Stickiness: Searching and Viewing Indexes
This figure replicates Figure 5 using the searching and viewing sub-indexes as alternative measures of retail investor attention. The specification is $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the cumulative return from the closing price on day $t-1$ to the closing price on day $t+h$. $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if stock i's attention stickiness exceeds the panel median on day t, and zero otherwise. We let h range from 0 to 40. The square-dotted line shows $\beta_{1,h} + \gamma_h$, capturing return dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows $\beta_{1,h} + \beta_{2,h} + \beta_{3,h} + \gamma_h$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. Panel (a) uses the searching index as the measure of abnormal retail attention; Panel (b) uses the viewing index. The sample period is from January 2018 to December 2024.

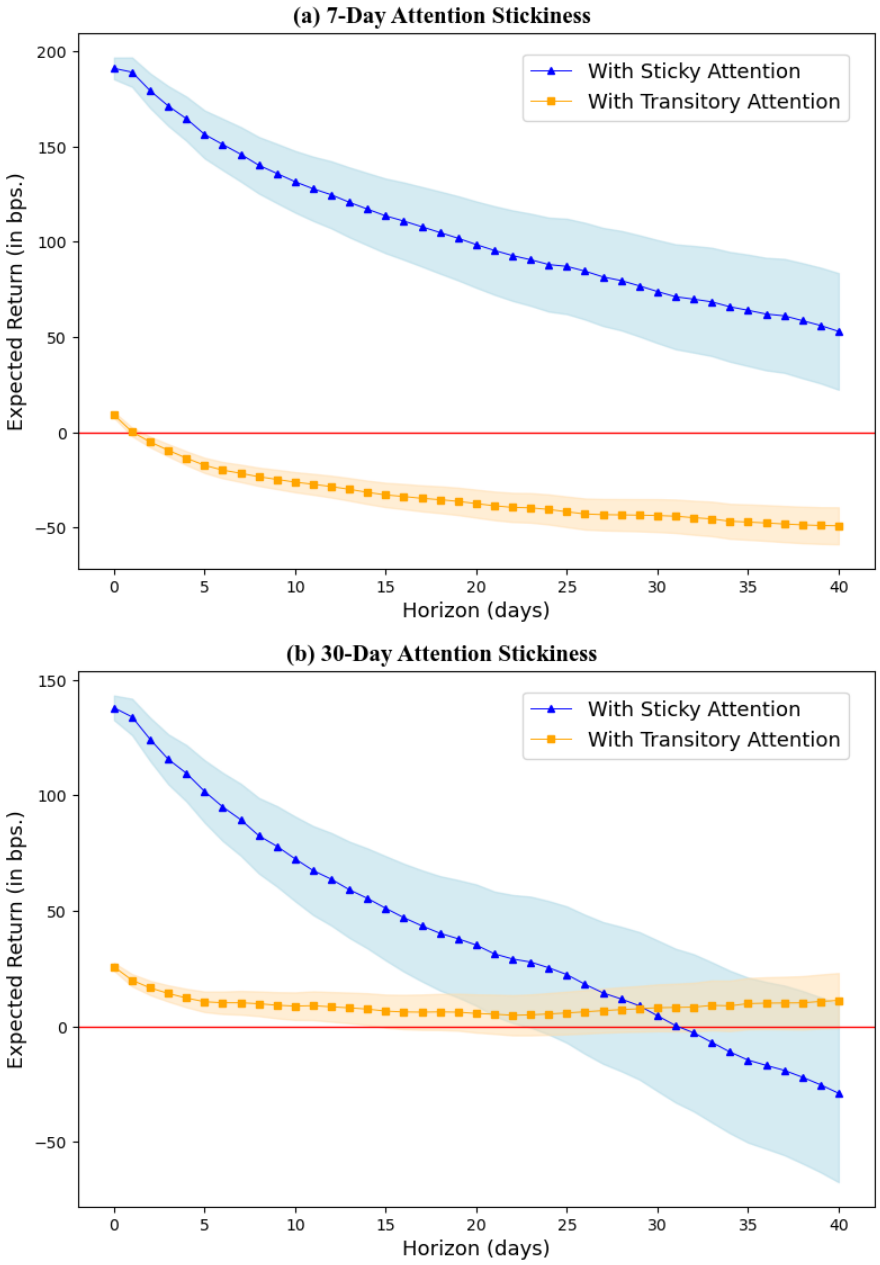


Figure A.4: Return Dynamics Conditional on Attention Stickiness: 7- and 30-Day Stickiness Horizons
This figure replicates Figure 5 using 7- and 30-day attention stickiness ratios in place of the baseline 1-day ratio. The specification is $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the cumulative return from the closing price on day $t-1$ to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention, and $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if the relevant stickiness ratio exceeds the panel median on day t, and zero otherwise. We let h range from 0 to 40. The square-dotted line shows $\beta_{1,h} + \gamma_h$, capturing return dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows $\beta_{1,h} + \beta_{2,h} + \beta_{3,h} + \gamma_h$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. Panel (a) classifies attention stickiness using the 7-day ratio ($STK_{i,t,7}$); Panel (b) uses the 30-day ratio ($STK_{i,t,30}$). The sample period is from January 2018 to December 2024.

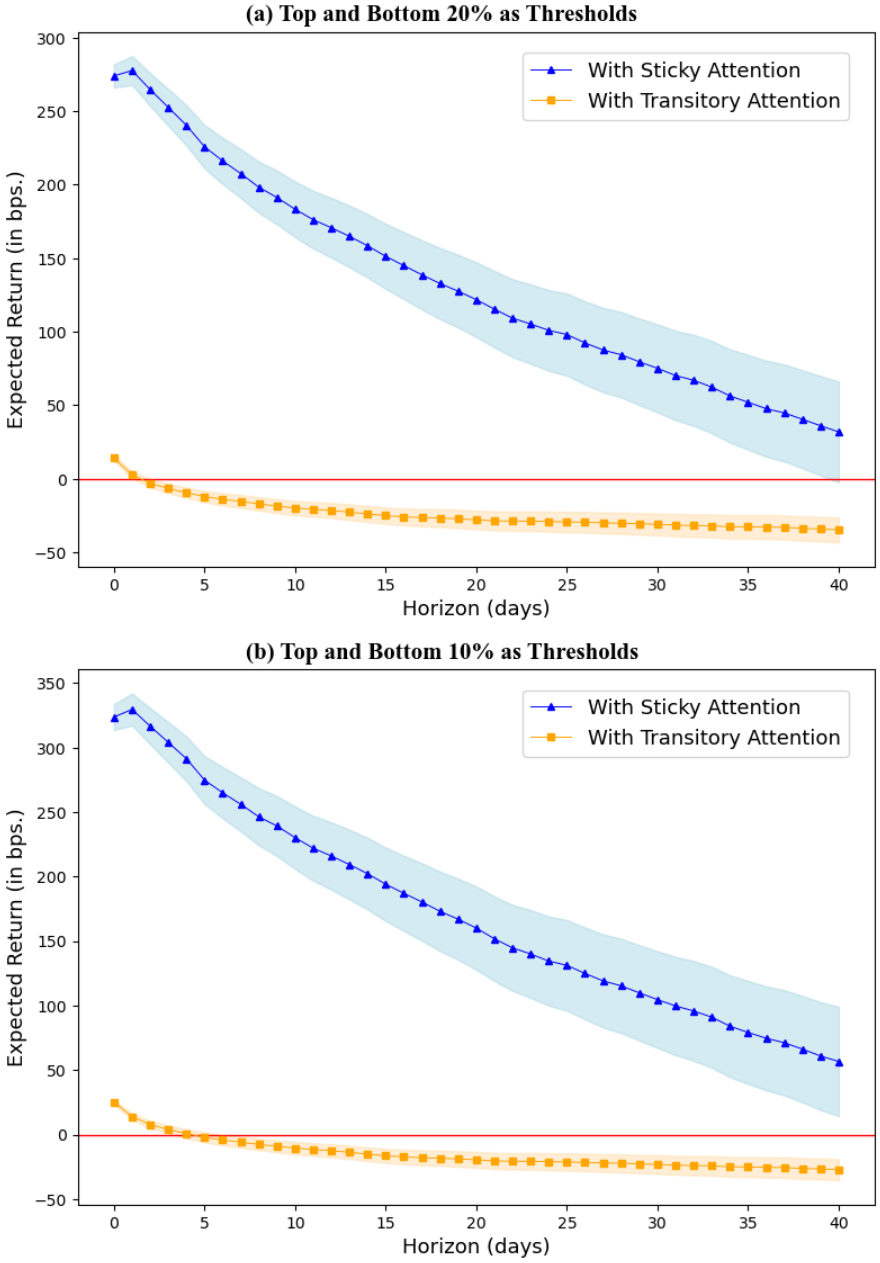


Figure A.5: Return Dynamics Conditional on Attention Stickiness: Alternative Classification Thresholds
This figure replicates Figure 5 using stricter thresholds to classify sticky and transitory attention. The specification is $r_{i,t\to t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times ATT_{i,t} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the cumulative return from the closing price on day $t-1$ to the closing price on day $t+h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention. $\mathbb{I}_{i,t}^{STK}$ equals one when stock i's attention stickiness ratio enters the top 20th (Panel a) or top 10th (Panel b) percentile of the cross-sectional distribution on day t, and zero when it falls below the corresponding bottom percentile; observations between the two thresholds are excluded. We let h range from 0 to 40. The square-dotted line shows $\beta_{1,h} + \gamma_h$, capturing return dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows $\beta_{1,h} + \beta_{2,h} + \beta_{3,h} + \gamma_h$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. The sample period is from January 2018 to December 2024.

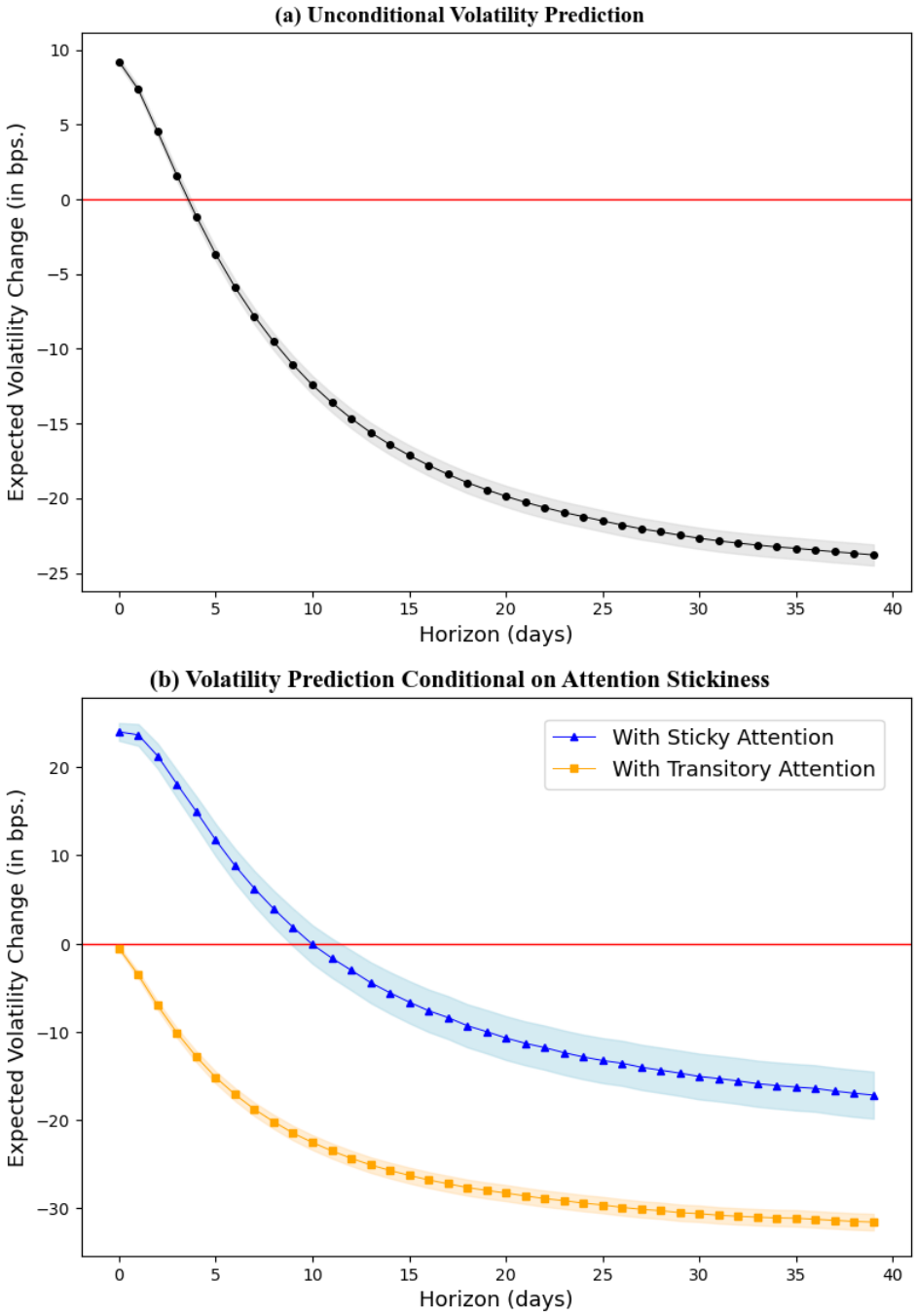


Figure A.6: Volatility Dynamics with Abnormal Attention and Attention Stickiness
This figure reports estimates from panel regressions of volatility dynamics: $\Delta Vol_{i,t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the change in GARCH(1,1) conditional volatility from day t to day $t + h$. $ATT_{i,t}$ is the standardized log-abnormal retail attention, and $STK_{i,t}$ denotes the standardized attention stickiness ratio. We let h range from 0 to 40. Panel (a) plots the unconditional estimates of $\beta_{1,h}$, representing a standard deviation shock to abnormal attention, obtained by setting $\beta_{2,h} = \beta_{3,h} = 0$ in the full specification. Panel (b) plots the estimates conditional on attention stickiness: the square-dotted line shows $\beta_{1,h}$, capturing volatility dynamics of a standard deviation shock to abnormal attention for transitory-attention stocks, and the triangle-dotted line shows $\beta_{1,h} + \beta_{2,h} + \beta_{3,h}$, capturing the effect for sticky-attention stocks. Shaded areas denote 95% confidence intervals; standard errors are clustered by firm and calendar day. The sample period is from January 2018 to December 2024.

Table A.1: Variable Descriptions
This table defines the variables used in the analysis. Panel A covers attention and stickiness measures from the RoyalFlush (Tong-Hua-Shun) platform; Panel B defines stock and firm fundamental variables.

Variables	Descriptions
Panel A: Attention and Stickiness Variables	
Attention Index	Composite attention measure: value-weighted average of the searching and viewing indexes, where viewing receives greater weight reflecting its higher raw frequency.
Searching Index	Number of searches of stock i on day t divided by the total number of searches across all stocks on day t (%).
Viewing Index	Number of page views of stock i on day t divided by the total number of page views across all stocks on day t (%).
Abnormal Attention ($ATT_{i,t}$)	Standardized log-ratio of the current-day attention index to its 30-day rolling median, following Da et al. (2011). Normalized to mean zero and unit standard deviation within the full sample.
Stickiness Ratio ($STK_{i,t,k}$)	Fraction of users who viewed stock i on day $t-k$ and returned to view it on day t, where $k \in \{1, 7, 30\}$ trading days. Formally, $STK_{i,t,k} =$ Users View$_{i,t-k \cap t}$/Users View$_{i,t-k}$.
Panel B: Stock and Fundamental Variables	
Leverage	Total liabilities divided by total assets.
Market Cap	Natural log of total market capitalization (CNY).
ROA	Net income divided by total assets.
Total Asset	Natural log of total assets (CNY).
Stock Price	Daily closing price (CNY).
Stock Return	Daily holding-period return based on closing prices.
Turnover	Daily trading volume divided by total shares outstanding (%).
Volatility	Conditional volatility estimated via GARCH(1,1).

Table A.2: Correlation Matrix of Key Variables
This table reports pairwise Pearson correlations among the main variables. ATT is the standardized log-abnormal retail attention index; Lagged ATT is its one-day lag. STK is the 1-day attention stickiness ratio; STK Indicator equals one if STK exceeds the panel median on that day; STK Orth is the orthogonalized stickiness measure obtained by retaining the intercept plus residual from regressing STK on attention, firm characteristics, and news indicators (see Table A.3). Volatility is estimated via a GARCH(1,1) model. Turnover is daily trading volume divided by shares outstanding (%). MarketCap is the natural log of total market capitalization (CNY). The sample period is from January 2018 to December 2024.

	Return	Lagged Return	ATT	Lagged ATT	STK	STK Dummy	STK Orth	Collection Change	Volatility	Turnover	MarketCap
Return	1										
Lagged Return	0.058	1									
ATT	0.276	0.219	1								
Lagged ATT	-0.012	0.270	0.821	1							
STK	0.376	-0.143	0.249	-0.089	1						
STK Dummy	0.231	-0.132	0.141	-0.087	0.773	1					
STK Orth	0.283	-0.001	0.089	-0.179	0.832	0.696	1				
Collection Change	0.475	0.136	0.358	0.117	0.282	0.127	0.157	1			
Volatility	0.015	0.074	0.333	0.415	0.116	0.074	-0.036	0.061	1		
Turnover	0.107	0.183	0.478	0.429	0.287	0.173	-0.035	0.160	0.576	1	
MarketCap	0.011	0.012	-0.012	-0.013	0.160	0.134	0.002	-0.006	-0.163	-0.177	1

Table A.3: Orthogonalization of Attention Stickiness
This table reports OLS estimates from the first-stage regression used to construct the orthogonalized attention stickiness measure. The specification is $STK_{i,t,k} = \alpha + \beta \times \mathbb{X}_{i,t} + \epsilon_{i,t}$, where the control matrix $\mathbb{X}_{i,t}$ includes: retail attetnion, its one-day lag, log market capitalization, turnover ratio, and indicators for positive, negative, and neutral news articles (from SuperMind). Columns (1) to (3) report results for the 1-, 7-, and 30-day attention stickiness ratios, respectively. The orthogonalized measure $STK^{\perp}_{i,t}$ is then defined as the estimated intercept $\hat{\alpha}$ plus the regression residual $\hat{\epsilon}_{i,t}$, ensuring orthogonality to all included controls by construction. The sample period is from January 2018 to December 2024.

	1-Day Attention Stickiness (1)	7-Day Attention Stickiness (2)	30-Day Attention Stickiness (3)
Constant	0.121***	-0.018***	0.105***
Attention	0.012***	0.008***	0.009***
Lagged Attention	-0.010***	-0.004***	-0.006***
Lagged Return	-0.701***	0.292***	0.476***
Market Value	0.015***	0.014***	0.013***
Turnover	0.896***	0.734***	0.557***
Bad News Indicator	0.003***	0.011***	0.010***
Good News Indicator	-0.007***	0.005***	0.003***
Neutral News Indicator	0.004***	0.013***	0.009***
$AdjR^2$	0.284	0.242	0.192
Observations	4,671,745	4,671,745	4,671,725

Table A.5: Cross-Sectional Heterogeneity

This table reports coefficient estimates from the return-predictability specification of Table 4, estimated separately for cross-sectional subgroups defined by sorting on stock characteristics. The specification is $r_{i,t\to t+10} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$, where the dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+10$. $ATT_{i,t}$ is the standardized log-abnormal retail attention index, and $STK_{i,t}$ denotes the standardized attention stickiness ratio. Panel A sorts on firm fundamentals: size (log market cap), illiquidity (Amihud ratio), and market-to-book ratio. Panel B sorts on trading characteristics: prior-month return, return volatility, and turnover. Panel C sorts on Chinese-market-specific conditions: short-sale eligibility (short-sale-eligible stocks face fewer limits to arbitrage, so mispricing should correct faster) and Hong Kong dual-listing (dual-listed stocks attract additional institutional scrutiny). Stocks are sorted into quintile groups and portfolios are rebalanced monthly; "High" ("Low") indicates estimates for the top (bottom) quintile. All regressions include firm and time fixed effects; standard errors are clustered by firm and calendar day. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

Panel A: Fundamentals

	Size			Illiquidity			Market-to-Book		
	Big	Small	Big - Small	High	Low	High - Low	High	Low	High - Low
ATT	-14.058***	-59.274***	45.216***	-47.858***	-18.080***	-29.778***	-34.286***	-23.952***	-10.335**
STK	-21.510***	-35.319***	13.809**	-46.171***	-21.922***	-24.249***	-30.950***	-27.385***	-3.565
$ATT \times STK$	5.811***	-3.857	9.668***	3.527	-0.896	4.423	5.720***	1.097	4.624*

Panel B: Trading Information

	Stock Return			Return Volatility			Turnover		
	High	Low	High - Low	High	Low	High - Low	High	Low	High - Low
ATT	-37.730***	-33.409***	-4.321	-48.273***	-25.235***	-23.038***	-52.887***	-27.065***	-25.822***
STK	-49.047***	-24.600***	-24.447***	-45.011***	-21.209***	-23.802***	-44.563***	-15.156***	-29.408***
$ATT \times STK$	5.275***	2.802*	2.473	3.939*	-1.564	5.503**	6.002***	0.254	5.748*

Panel C: Chinese Market Special Regulations

	Short-sale Permission			Hong Kong Dual-listing		
	Yes	No	Yes - No	Yes	No	Yes - No
ATT	-15.915***	-43.294***	27.379***	-6.483	-31.113***	24.630***
STK	-28.737***	-37.294***	8.556***	-27.407***	-34.094***	6.687
$ATT \times STK$	3.869***	0.757	3.112*	4.326	2.361**	1.965

Table A.6: Return Predictability: Attention Indicator Interacted with Attention Stickiness
This table reports coefficient estimates of return predictability using indicator versions of both attention and stickiness. The specification is $r_{i,t+1\to t+h} = \beta_{1,h} \times \mathbb{I}_{i,t}^{ATT} + \beta_{2,h} \times \mathbb{I}_{i,t}^{STK} + \beta_{3,h} \times \mathbb{I}_{i,t}^{ATT} \times \mathbb{I}_{i,t}^{STK} + \gamma_h \times r_{i,t-1} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+h$. $\mathbb{I}_{i,t}^{ATT}$ is an indicator equal to one if stock i's abnormal attention ($ATT_{i,t}$) exceeds the panel median on day t; $\mathbb{I}_{i,t}^{STK}$ is an indicator equal to one if stock i's 1-day attention stickiness ratio exceeds the panel median on day t. This robustness specification verifies that the baseline results of Table 4 are not sensitive to the use of continuous $ATT_{i,t}$. All regressions include firm and time fixed effects; standard errors are clustered by firm and calendar day. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Same-day	Future Horizons (days)				
	$h=0$	$h=1$	$h=5$	$h=10$	$h=20$	$h=40$
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{I}^{ATT}$	6.183***	-8.309***	-17.144***	-22.830***	-35.571***	-54.964***
	(1.145)	(0.924)	(2.350)	(3.285)	(4.698)	(6.512)
$\mathbb{I}^{STK}$	51.730***	-5.177***	-30.568***	-56.838***	-98.388***	-170.003***
	(0.938)	(0.787)	(1.995)	(2.943)	(4.684)	(7.215)
$\mathbb{I}^{ATT} \times \mathbb{I}^{STK}$	145.585***	4.464***	-8.441***	-13.731***	-16.346***	-10.375
	(1.716)	(1.015)	(2.335)	(3.311)	(4.752)	(6.414)
γ	0.067***	-0.000	-0.044***	-0.076***	-0.125***	-0.178***
	(0.004)	(0.004)	(0.009)	(0.011)	(0.015)	(0.016)
$AdjR^2$	0.330	0.248	0.256	0.247	0.247	0.257
Observations	4,671,745	4,668,620	4,656,121	4,640,501	4,609,177	4,546,441
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time

Table A.7: Conditional Return Prediction with Alternative Abnormal Attention Settings
This table reports coefficient estimates from panel regressions of future stock returns on abnormal retail attention, the attention stickiness indicator, and their interaction, with alternative settings of abnormal attention. The specification is $r_{i,t+1\to t+h} = \beta_{1,h}\times ATT_{i,t}+\beta_{2,h}\times STK_{i,t}+\beta_{3,h}\times ATT_{i,t}\times STK_{i,t}+\gamma_h\times r_{i,t-1}+\alpha_{i,h}+\delta_{t,h}+\epsilon_{i,t,h}$. The dependent variable is the holding-period return from the closing price on day t to the closing price on day $t+h$; Column (1) reports the contemporaneous association, where the dependent variable is the close-to-close return on day t. Independent variables include normalized abnormal attention, attention stickiness, and their interaction. One-day lagged stock return is added as a control variable. Panel A produces abnormal attention as the log deviation of the daily attention index from its own 8-week (40-trading-day) median. Panels B and C uses 4-week (20-trading-day) and 2-week (10-trading-day) as the benchmarks. All regressions include firm and time fixed effects; standard errors are clustered by firm and time. Coefficients are in basis points. $\beta_{1,h}$, $\beta_{2,h}$, and $\beta_{3,h}$ are reported in the table. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Same-day	Future Horizons (days)				
	$h=0$	$h=1$	$h=5$	$h=10$	$h=20$	$h=40$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Abnormal Attention with an 8-week Benchmark						
ATT	25.644***	-8.958***	-24.459***	-33.716***	-42.619***	-45.588***
	(0.811)	(0.719)	(1.723)	(2.328)	(3.260)	(4.638)
STK	75.430***	-1.356**	-16.336***	-30.798***	-58.085***	-107.502***
	(0.923)	(0.550)	(1.361)	(1.925)	(3.024)	(4.692)
$ATT\times STK$	41.638***	6.825***	3.416***	0.694	1.280	8.556***
	(0.422)	(0.311)	(0.800)	(1.095)	(1.472)	(1.852)
Adj. R^2	0.448	0.249	0.257	0.248	0.249	0.259
Observations	4,640,361	4,637,237	4,624,741	4,609,121	4,577,804	4,515,087
Panel B: Abnormal Attention with a 4-week Benchmark						
ATT	39.476***	-7.665***	-19.418***	-24.765***	-28.697***	-29.442***
	(0.812)	(0.655)	(1.580)	(2.101)	(2.788)	(3.677)
STK	71.223***	-2.872***	-21.412***	-38.553***	-69.030***	-119.711***
	(0.893)	(0.557)	(1.380)	(1.982)	(3.047)	(4.664)
$ATT\times STK$	40.101***	7.749***	6.640***	5.539***	7.527***	15.376***
	(0.402)	(0.292)	(0.746)	(1.037)	(1.353)	(1.661)
Adj. R^2	0.461	0.251	0.260	0.249	0.248	0.258
Observations	4,703,144	4,700,019	4,687,519	4,671,894	4,640,557	4,577,804
Panel C: Abnormal Attention with a 2-week Benchmark						
ATT	55.534***	-5.841***	-13.259***	-14.222***	-14.419***	-8.766***
	(0.861)	(0.634)	(1.548)	(1.955)	(2.536)	(3.101)
STK	66.608***	-4.110***	-25.692***	-45.533***	-78.248***	-130.671***
	(0.857)	(0.576)	(1.422)	(2.040)	(3.107)	(4.674)
$ATT\times STK$	38.430***	8.279***	9.207***	9.510***	12.556***	19.904***
	(0.377)	(0.275)	(0.653)	(0.879)	(1.156)	(1.454)
Adj. R^2	0.480	0.250	0.260	0.252	0.249	0.258
Observations	4,734,573	4,731,447	4,718,943	4,703,312	4,671,960	4,609,177
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time

Table A.8: Volatility Predictability Conditional on Attention Stickiness
This table reports coefficient estimates from panel regressions of future volatility changes on abnormal retail attention, attention stickiness, and their interaction. The specification is $\Delta Vol_{i,t+h} = \beta_{1,h} \times ATT_{i,t} + \beta_{2,h} \times STK_{i,t} + \beta_{3,h} \times ATT_{i,t} \times STK_{i,t} + \alpha_{i,h} + \delta_{t,h} + \epsilon_{i,t,h}$. The dependent variable is the change in GARCH(1,1) conditional volatility from day t to day $t+h$. $ATT_{i,t}$ is the standardized abnormal retail attention, and $STK_{i,t}$ denotes the standardized attention stickiness ratio. Panel A establishes the unconditional volatility response to retail attention, obtained by setting $\beta_{2,h} = \beta_{3,h} = 0$ in the full specification. Panel B tests whether sticky and transitory attention generate asymmetric volatility dynamics: the coefficient $\beta_{1,h}$ captures the effect for transitory-attention stocks and $\beta_{1,h} + \beta_{3,h}$ captures the effect for sticky-attention stocks. All regressions include firm and time fixed effects; standard errors are clustered by firm and day. Coefficients are in basis points. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. The sample period is from January 2018 to December 2024.

	Future Horizons (days)				
	$h = 1$	$h = 5$	$h = 10$	$h = 20$	$h = 40$
	(1)	(2)	(3)	(4)	(5)
Panel A: Unconditional Prediction					
ATT	9.229***	-1.139***	-11.065***	-19.427***	-23.793***
	(0.117)	(0.200)	(0.297)	(0.356)	(0.358)
Adj. R^2	0.195	0.161	0.187	0.224	0.252
Panel B: Conditional on Attention Stickiness					
ATT	3.713***	-6.788***	-15.595***	-22.783***	-26.552***
	(0.103)	(0.242)	(0.336)	(0.386)	(0.392)
STK	6.118***	6.173***	5.044***	3.568***	1.523***
	(0.128)	(0.179)	(0.217)	(0.267)	(0.315)
$ATT \times STK$	6.567***	6.767***	5.380***	4.055***	3.962***
	(0.081)	(0.089)	(0.103)	(0.120)	(0.127)
Adj. R^2	0.327	0.193	0.200	0.230	0.255
Observations	4,668,620	4,656,121	4,640,501	4,609,177	4,546,441
FE	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time
Clusters	Firm/Time	Firm/Time	Firm/Time	Firm/Time	Firm/Time