Watch and Foresee: Evaluate Video Media Sentiment through Large Language Model

Fuwei Jiang

School of Economics, Xiamen University, Xiamen 361005, China;

Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361005, China;

jfuwei@gmail.com

Fujing Jin

School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China;

fujing_jin@163.com

Tian Ma

School of Economics, Minzu University of China, Beijing 100081, China;

mark8938@qq.com

Watch and Foresee: Evaluate Video Media Sentiment through Large Language Model

Abstract

We study whether sentiment conveyed through videos predicts stock returns. Using a large language model, we construct a daily market-level visual sentiment index (SVS) from facial expressions in short videos posted by influencers on Douyin and compare it with a caption-based textual sentiment measure (STS). SVS predicts future market returns through an initial positive response followed by partial reversal, with the predictability concentrated in intraday rather than overnight returns. We further show that SVS contains predictive information distinct from STS, and textual sentiment dampens the initial positive price response associated with visual sentiment. Additional tests support a sentiment-based interpretation of SVS: its predictive effects are more pronounced in upward markets, retail-oriented portfolios and active trading environments, while disagreement between SVS and STS also predicts trading volume. Overall, short-video visual sentiment contains market-relevant information beyond accompanying texts and helps explain the timing and dynamics of sentiment-driven price responses.

Keywords: Large language model, Video-based sentiment, Return predictability

1. Introduction

Investor sentiment remains a central topic in the return predictability literature. Early studies measure sentiment using either structured proxies from market data (Baker & Wurgler 2006; Huang *et al.* 2015) or tone extracted from text (Jiang *et al.* 2019; Chen *et al.* 2023). More recent evidence shows that sentiment is also embedded in images, including news photographs (Obaid & Pukthuanthong 2022). Yet an increasing share of information is now created and disseminated through videos on social media. Relative to text or static images, videos convey emotion through richer and more dynamic cues, especially facial expressions, which may affect investor perceptions and prices in ways not fully captured by accompanying words alone (Blankespoor *et al.* 2023). We therefore examine whether sentiment embedded in videos predicts equity prices.

However, quantifying sentiment in videos poses a first-order challenge: video data are unstructured, high-dimensional, and costly to process at scale. We address these constraints by focusing on short videos, which convey information in a concentrated format and are more computationally tractable than long-form content. Our sample comes from Douyin, the largest short-video platform in China, which permits systematic collection of video content. China also provides a natural setting because retail investors account for a substantial share of trading and are more exposed to sentiment and attention frictions than institutional investors. These features make China well suited for our empirical analysis.

To study how short videos affect equity prices, we focus on content with broad dissemination potential. Following recent work on social media visibility and market responses (Jagolinzer 2024; Merkley *et al.* 2024), we restrict attention to videos posted by influential creators. Specifically, we sample videos from the top 50 accounts by follower count, increasing the likelihood that sentiment shocks are observed by

investors while reducing noise from low-exposure content. We then construct a daily short-video visual sentiment index (SVS) from facial expressions in these videos. Following Chang *et al.* (2024), we use *OpenFace* to extract frame-level facial action measures and classify frames as reflecting positive or negative emotion. We aggregate these signals to the video level and then form a follower-weighted daily index across videos.

We begin by examining whether visual sentiment extracted from short videos predicts market returns. SVS predicts returns at short horizons through a positive first-lag response followed by partial reversal over subsequent trading days. By contrast, textual sentiment (STS) extracted from the same videos exhibits a sharper and more immediate reversal pattern. Although SVS and STS are extracted from the same underlying videos, they differ in modality and therefore need not be incorporated into prices in the same way. STS captures verbalized affect, which is easier to parse and more readily translated into an explicit directional signal. SVS, by contrast, captures nonverbal affective cues conveyed through dynamic facial expressions, which are richer, more socially salient, and potentially more affective but less codified. This difference in information processing helps explain why STS exhibits a sharper reversal pattern, whereas SVS is associated with an initial positive response followed by partial reversal. The negative interaction between the two measures further suggests that they operate as distinct but cross-checking sentiment signals, with each helping to discipline the market response to the other.

For both measures, predictability is concentrated in intraday rather than overnight returns, indicating that sentiment embedded in short videos is incorporated primarily during trading hours. When SVS and STS enter jointly, both retain predictive power, and their interaction suggests that they dampen rather than reinforce each other's mispricing effects. Under the standard sentiment-pricing view that elevated sentiment initially pushes prices above fundamentals and is then followed by downward correction, SVS

appears to strengthen the subsequent reversal associated with STS, whereas STS attenuates the initial positive price response associated with SVS. Hence, these results indicate that SVS and STS contain distinct sentiment information.

To interpret the content of SVS, we show that its predictive effects are more structured in market up states, in retail-oriented portfolios, and in active trading environments—settings in which sentiment-driven trading is more likely to move prices and patterns that are difficult to reconcile with a purely risk-based explanation. We further show that disagreement between SVS and STS has implications for trading activity. Because the two measures are extracted from different modalities of the same underlying videos, their divergence captures heterogeneous interpretations rather than differences in dissemination. This disagreement raises trading volume in high-retail portfolios but lowers it in low-retail portfolios, consistent with disagreement-driven trading in retail-intensive settings and with conflicting signals being treated more as noise when retail participation is weaker.

Additional analyses reinforce the distinct role of SVS. Sentiment extracted from institutional videos is qualitatively similar but quantitatively weaker, while the predictive pattern of SVS remains intact after controlling for institutional video sentiment. Out-of-sample tests likewise show that SVS and STS contain different predictive content: STS sometimes performs better in statistical forecast accuracy, but SVS often delivers greater economic value, especially for the aggregate market and the Shanghai Stock Exchange market. Finally, an alternative measure constructed from only the first 30 seconds of each video yields weaker and less stable predictability than the baseline full-video SVS, suggesting that the sentiment relevant for prices unfolds over the course of the full video rather than being concentrated solely in the opening segment.

This study makes three contributions. First, we contribute to the sentiment and return-predictability literature by introducing a new daily sentiment measure based on short-form videos. Existing sentiment measures are largely built from structured market proxies (Baker & Wurgler 2006; Huang *et al.* 2015; Wang *et al.* 2022), surveys (Brown & Cliff 2005; Huang *et al.* 2022; Tiwari *et al.* 2022), and text-based signals (Da *et al.* 2014; Behrendt & Schmidt 2018; Jiang *et al.* 2019; Chen *et al.* 2023; Hirshleifer *et al.* 2025), with recent work incorporating still images (Obaid & Pukthuanthong 2022).

We move beyond these approaches by constructing a daily sentiment index from facial expressions in videos. Relative to text- and proxy-based measures, our index captures fine-grained affective variation from subtle facial cues and is less sensitive to wording and writing style, which broadens the set of settings in which sentiment can be measured. Using this measure, we show that video-based sentiment predicts aggregate returns over short horizons with a continuation-then-reversal pattern that concentrates in intraday, rather than overnight, returns—evidence that is difficult to reconcile with risk compensation or fundamental news and supports a sentiment interpretation of the index.

Second, we contribute to the growing empirical literature that uses unstructured video data as a source of economic information by explicitly exploiting the multimodal nature of short videos. Prior work leverages the presence of video content (Nekrasov *et al.* 2022; Cai *et al.* 2023; Crowley *et al.* 2025), extracts verbal information from videos (Gómez-Cram & Grotteria 2022), or studies video-based communication in both archival settings and experiments (Elliott *et al.* 2012; Andrade *et al.* 2015; Cade *et al.* 2020; Cronin *et al.* 2021). Related studies often rely on laboratory software that reads coarse emotional categories or on human ratings of facial traits and credibility from video stills and roadshow presentations (Blankespoor *et al.* 2017; Breaban & Noussair 2018; Huang *et al.* 2023).

Building on Chang *et al.* (2024), we use *OpenFace* 2.0 to extract frame-level facial-action measures and construct a visual sentiment index. Importantly, we go beyond a single dimension and treat short videos as multimodal objects. We extract visual and textual sentiment from the same videos, compare them, and study the economic content of their disagreement. This perspective differs from Hu and Ma (2025), who emphasize multiple channels but summarize detailed measures by extracting a common latent component; instead, we focus on the cross-modal gap and show that it is economically informative. Unlike Chang *et al.* (2024), who use facial cues in application videos to assess borrower creditworthiness, we study facial-expression-based sentiment in social-media videos and its market implications.

Third, our setting speaks to the social-network view of information diffusion on social media. Prior work formalizes influencer marketing as a network-seeding and scheduling problem—highlighting peer spillovers, multiple exposures, and forgetting in shaping engagement (Mallipeddi *et al.* 2022)—and shows that follower size impact engagement (Wies *et al.* 2023). In finance, related evidence and theory emphasize that influencer-originated messages can move prices through rapid diffusion and attention, with effects that tend to be short-lived and partially reversed as beliefs update (Pedersen 2022; Merkley *et al.* 2024), and discussions of social-media information transmission stress the role of highly connected nodes in accelerating dissemination (Jagolinzer 2024).

Against this backdrop, we bring an influencer-centered diffusion perspective into equity return predictability using short-video content. By focusing on influencer-generated videos—highly connected nodes in the platform network—and weighting content by follower reach, we operationalize dissemination intensity in a transparent way and study sentiment shocks that are more likely to be widely observed and priced.

The rest of the paper is organized as follows. Section 2 describes the data, sample, and variable construction. Section 3 examines the pricing effects of sentiment embedded in the visual and textual components of short videos. Section 4 provides a set of validation tests for short-video visual sentiment and assesses whether it captures sentiment-related information rather than alternative concerns. Section 5 presents further analyses, including the role of institutional video sentiment, out-of-sample performance, and an alternative measure based on the first 30 seconds of each video.

2. Data and sample

2.1 Data sources

Our primary data come from Douyin, China's version of TikTok, and the China Stock Market & Accounting Research (CSMAR) database. The sample spans January 1, 2022 to December 31, 2025. From Douyin, we collect short videos posted by personal finance influencers, together with video-level metadata such as posting time and engagement measures. We focus on videos posted by the top 50 personal finance influencers ranked by follower counts, while allowing the set of top accounts to evolve over time. If an influencer does not post any videos for an entire month or becomes unavailable on the platform, for example because of account suspension, we replace that account with an influencer of comparable follower base and reach. Applying this rolling replacement rule yields videos from 61 distinct influencers over the sample period, suggesting that turnover among top accounts is not particularly frequent and that many influencers remain both prominent and active over time. These influencers post frequently, with an average of about 19 active videos per day. As of December 31, 2025, the average follower count across the 61 influencers is 0.494 million, and the largest account has 2.243 million followers. Taken together, these features suggest that the selected videos are posted by accounts with substantial reach and are

therefore likely to have nontrivial influence on investor belief as well as the financial market. To ensure sufficient investor exposure while mitigating potential look-ahead concerns, we restrict attention to videos posted before 10:00 p.m. each day.

We clean the Douyin sample by removing duplicate uploads, reposts, and non-finance videos identified through keyword filters and manual checks. We also exclude videos with missing posting timestamps or missing engagement statistics. For the visual sentiment measures, we require that a video contain at least one detectable face for a nontrivial fraction of frames; videos with no face detections or severely corrupted frames are discarded.

We obtain stock-market data from CSMAR, including the returns for the aggregate market and individual stocks, Shanghai Stock Exchange (SSE) Composite Index, Shenzhen Stock Exchange (SZSE) Component Index, closing prices, and firm-level characteristics used for portfolio formation and control variables. We apply standard filters by excluding observations with missing returns or trading variables and by dropping stocks subject to special treatment or other trading irregularities. We merge the Douyin and CSMAR data using the posting date and the Chinese trading calendar to align social-media information with market outcomes.

2.2 Main variables

Short-video visual sentiment (SVS). We construct a daily market-level short-video visual sentiment index (SVS) from facial expressions in influencer videos. We begin by extracting frames from each Douyin short video and applying *OpenFace 2.0*, an open-source facial behavior measurement system that performs face detection and landmark localization and outputs frame-level Facial Action Coding System (FACS)-based action units (AUs). OpenFace reports the intensity or presence of 18 AUs, each

corresponding to the activation of specific facial muscles, such as cheek raising, brow movement, and lip-corner movement.

Following Chang *et al.* (2024), we map AU configurations into discrete emotional states and focus on happiness and fear. Specifically, we classify a frame as *happy* when its AU configuration is consistent with a smile, captured by the joint activation of AU6 (cheek raiser) and AU12 (lip corner puller). We classify a frame as *fearful* when its AU configuration reflects fear, based on the joint activation of AU1 (inner brow raiser), AU2 (outer brow raiser), AU4 (brow lowerer), AU5 (upper lid raiser), AU7 (lid tightener), AU20 (lip stretcher), and AU26 (jaw drop).

For each video v, we count the number of valid frames classified as *happiness* and *fear*, denoted H_v and F_v, where valid frames are those with successful face detection and usable AU outputs. We then define video-level net visual sentiment as:

$$SVS_{v,t} = \frac{H_v - F_v}{N_v}, \quad (1)$$

where N_v is the number of valid frames in the video v. This construction yields a bounded and interpretable measure that increases when *happy* expressions dominate *fearful* expressions and decreases when *fearful* expressions dominate.

Finally, we aggregate video-level sentiment into a daily market-level index based on the video posting date. Let w_v denote the follower count of the influencer posting video v. For each calendar day *t*, we first compute a calendar-day sentiment measure as the follower-weighted average of video-level sentiment across all videos posted on that day:

$$SVS_t^c = \frac{\sum_{v \in t} w_v \cdot SVS_v}{\sum_{v \in t} w_v}. \quad (2)$$

Weighting by follower counts assigns greater influence to content with higher dissemination potential and aligns the daily index with the potential reach of the underlying sentiment signal.

Because stock markets are closed on weekends and holidays whereas SVS_t^c is constructed for every calendar day, we convert the calendar-day measure into a trading-day measure by averaging sentiment over non-trading days that fall immediately before a trading day. Specifically, if the market is open on trading day *t* but was closed for the preceding *k* calendar days, we assign the following trading-day sentiment measure:

$$SVS_t = \frac{1}{k+1}\sum_{j=0}^{k} SVS_{t-j}^c. \quad (3)$$

Thus, when no non-trading day intervenes, $SVS_t = SVS_t^c$; otherwise, the trading-day measure averages sentiment accumulated over the intervening non-trading days and day *t* itself.

Short-video textual sentiment (STS). We construct a daily textual sentiment index (STS) from both the written captions and the spoken words in short videos. For each video, we first collect the caption text provided on Douyin. We then extract the audio track and use *Librosa* to standardize the audio input, including resampling and segmentation, and to exclude clips with unusable audio quality. Next, we apply an automatic speech recognition (ASR) model to transcribe the spoken content into text. We concatenate the caption and the transcript to form a unified video-level text corpus and then apply a machine-learning sentiment classifier to obtain a video-level textual sentiment score, denoted by STS_v.

We aggregate video-level textual sentiment into a daily market-level index using the same follower-count weights and calendar-to-trading-day alignment as in the construction of SVS. Specifically, if trading day *t* is preceded by *k* non-trading calendar days, we define

$$STS_t = \frac{1}{k+1}\sum_{j=0}^{k} STS_{t-j}^c, STS_t^c = \frac{\sum_{v\in t} w_v \cdot STS_v}{\sum_{v\in t} w_v}. \quad (4)$$

This matched aggregation ensures that differences between SVS and STS reflect differences between the visual and textual modalities rather than differences in sample composition or weighting schemes.

Market portfolio excess return (MKT). The market portfolio excess return is the daily excess return on the value-weighted stock market portfolio, computed as the market portfolio return minus the risk-free rate.

Excess return for the SSE market (SH index). $SH\ index$ is the daily excess return on the SSE Composite Index.

Excess return for the SZSE market (SZ index). $SZ\ index$ is the daily excess return on the SZSE Component Index.

Intraday return (intraday). Intraday return is measured from the market open to the market close within the same trading day. Specifically, intraday return for day *t* is defined as the percentage change from the opening price to the closing price on day *t*.

Overnight return (overnight). Overnight return is measured from the previous day's close to the next trading day's open. Specifically, overnight return for day *t* is defined as the percentage change from the closing price on the last trading day to the opening price on day *t*.

Trading volume (TV). Market-wide trading volume is measured as the aggregate trading value across all stocks in the market.

Day-of-the-week dummies. To account for calendar regularities, we include indicator variables for Tuesday through Friday, with Monday omitted as the benchmark category.

Other control variables. In the predictive regressions, we also control for standard return dynamics by including lagged market returns and lagged squared market returns, following the common practice in short-horizon return predictability tests.

2.3 Summary statistics

Table 1 reports summary statistics for the sentiment measures and market returns. Panel A shows that both sentiment measures are available for 1,212 daily observations. SVS has a positive mean of 0.061, whereas STS has a lower positive mean of 0.021. STS exhibits greater overall dispersion than SVS, with a standard deviation of 0.054 versus 0.039. The mean and median are close for both measures.

Panel B reports the Pearson correlation between SVS and STS. The correlation is −0.070 and statistically significant at 1% level, although economically small. This result suggests that video-based and text-based sentiment capture distinct, and potentially conflicting, aspects of investor sentiment rather than redundant information.

Panel A also shows that the mean excess market return (MKT) is positive at 0.078, while average returns for the SSE composite index and SZSE component index are close to zero. For both exchanges, intraday returns are positive on average, whereas overnight returns are negative. Intraday returns are also more volatile than overnight returns, especially in the Shenzhen Stock Exchange. These patterns suggest that a substantial share of return variation is concentrated during trading hours.

[Insert Table 1 about here]

3. Sentiment embedded in videos

3.1 Visual sentiment and market returns

We first examine the predictive power of short-video sentiment for future aggregate stock market returns using the following predictive regression:

$$R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \gamma X_t + \varepsilon_t, \quad (5)$$

where R_t denotes market returns on trading day *t*, including MKT, $SH\ index$, and $SZ\ index$. SVS represents the standardized short-video sentiment index constructed using a large language model (see

Section 2.2 for details). X is a spectrum of control variables, including five lags of both market returns and square of market returns, as well as day-of-the-week dummies, excluding Monday.

Table 2 reports the results. Column (1) corresponds to MKT, while columns (2) and (3) correspond to $SH\ index$, and $SZ\ index$, respectively. Across all three specifications, lagged SVS exhibits economically meaningful short-horizon predictive power. For example, a one-standard-deviation increase in SVS_{t-1} is associated with a 21.5-basis-point increase in MKT_t, indicating that higher short-video visual sentiment predicts higher next-day market returns. By contrast, the coefficient on SVS_{t-2} is negative and statistically significant for MKT, and $SH\ index$, suggesting a reversal effect at slightly longer horizons.

To assess the persistence of this reversal, we conduct joint significance tests across lagged SVS coefficients. The sum of coefficients from the first lag onward is positive and statistically significant across all three return series. By contrast, the sum of coefficients from the second through the fifth lag is negative and statistically significant for MKT, and $SH\ index$, and marginally significant for $SZ\ index$. This pattern indicates an initial positive return response followed by partial reversal over subsequent days, and the reversal does not fully offset the initial return response within five trading days.

Adjusted R^2 values are modest but comparable across specifications, which is typical for daily return predictability regressions. Overall, the evidence suggests that SVS contains incremental information for forecasting short-term market returns, with clear short-run positive drift followed by partial reversal, even after controlling for standard return dynamics and calendar effects.

[Insert Table 2 about here]

3.2 Sentiment and return decomposition

We further decompose daily returns on day *t* into overnight and intraday components, where overnight returns are computed from the closing price on the previous trading day to the opening price on day *t*, whereas intraday returns are computed from the opening to the closing price on day *t*. We use this decomposition to examine when sentiment is incorporated into prices.

As reported in Panel A of Table 3, SVS shows much stronger predictive power for intraday returns than for overnight returns. For both the Shanghai and Shenzhen Stock Exchange, the coefficient on SVS_{t-1} is positive and highly significant for intraday returns, indicating that higher short-video sentiment predicts higher next-day trading-session returns. By contrast, SVS has little predictive power for overnight returns. This result suggests that the information in SVS is incorporated primarily during trading hours rather than overnight.

The subsequent intraday coefficients indicate partial reversal over the following days. In particular, the sum of coefficients from the second through the fifth lag is negative and statistically significant for both SSE and SZSE intraday returns, whereas no comparable joint significance is detected for overnight returns. Taken together, these results indicate that SVS generates an initial positive intraday price response followed by partial reversal within subsequent trading days. Overall, SVS-based predictability is concentrated during the trading day and largely absent from overnight returns.

[Insert Table 3 about here]

To further examine the source of short-video return predictability, we distinguish between the overall level of sentiment conveyed by a video and the extent to which sentiment fluctuates within the video itself. While SVS captures the average sentiment embedded in short videos, videos are inherently dynamic and may differ substantially in how sentiment evolves over time. A video with a given sentiment level may convey that sentiment in a smooth and stable manner or through sharp emotional swings. Such within-

video variation may matter for return predictability if investors respond not only to the level of sentiment, but also to the intensity or instability of sentiment as it unfolds over the course of the video.

To capture this dynamic feature, we construct a measure of video sentiment innovation volatility, denoted by FS, as the standard deviation of the first difference in second-by-second video sentiment. We then augment the predictive regression with lagged FS, and the interaction between SVS and FS, in order to examine whether the predictive effect of short-video sentiment depends on the volatility of within-video sentiment innovations.

Table 4 shows that the coefficient on SVS_{t-1} remains positive and statistically significant for MKT, $SH\ index$, and $SZ\ index$ after controlling for lagged FS, indicating that the baseline predictive power of short-video sentiment is robust. By contrast, FS_{t-1} has little direct predictive power for aggregate market returns, although it is weakly positively related to overnight returns. In addition, the interaction term between SVS and FS is generally insignificant, with only weak evidence for $SZ\ index$. The SZSE exhibits some heterogeneity in this regard, possibly because its investor base contains a larger share of emotionally driven retail traders[1]. Overall, these results suggest that the predictive content of short videos is driven mainly by sentiment levels rather than by high-frequency fluctuations in sentiment within the video.

[Insert Table 4 about here]

3.3 Visual vs textual sentiment

Although short videos represent a new medium for information dissemination, they add little value unless they convey information beyond text (Goldstein *et al.* 2024). We therefore benchmark video-based

[1] In the portfolio analysis in Section 4.2, we proxy retail intensity using stock price. The resulting high-retail portfolio contains 59.22% Shenzhen-listed stocks, which is consistent with, though not by itself proof of, greater retail investor participation in that market.

sentiment (SVS) against caption-based textual sentiment (STS) and begin by estimating the following predictive regression for STS:

$$R_t = \alpha + \sum_{i=1}^{5} \beta_i STS_{t-i} + \gamma X_t + \varepsilon_t, \quad (6)$$

where STS denotes the standardized sentiment embedded in short video captions using machine learning (see Section 2.2 for details).

Table 5 reports time-series regression results for STS. In contrast to short-video visual sentiment, STS exhibits a sharper but less persistent predictive pattern. The coefficient on STS_{t-1} is negative and statistically significant across all three market returns, indicating that more positive textual sentiment predicts lower next-day returns. The coefficient on STS_{t-2} is positive, but statistically significant only for $SZ\ index$. This pattern suggests an immediate negative return response followed by only partial correction. Higher-order lags are generally small, although STS_{t-5} is positive and statistically significant across all three market returns.

The joint significance tests further highlight the differences between STS and SVS. The sum of coefficients from the first through the fifth lag is negative and statistically significant in all three specifications. By contrast, the sum of coefficients from the second through the fifth lag is small and statistically insignificant throughout. This pattern indicates that the strong initial negative effect of STS is not offset by subsequent lags within five trading days.

Compared with Table 2, STS displays a more abrupt initial response and less orderly subsequent adjustment. SVS generates an initial positive return response followed by partial reversal, whereas STS begins with a strong negative response and exhibits weaker subsequent correction. These results suggest that the market reacts more quickly to textual sentiment, and the initial response is followed by a next-day reversal that is consistent with the correction of sentiment-driven overreaction documented in the return

predictability literature (Huang *et al.* 2015; Edmans *et al.* 2022). By contrast, video-based sentiment appears to be incorporated more gradually, possibly because its emotional content is less readily captured by standard market processing.

[Insert Table 5 about here]

Secondly, we examine the predictive pattern after return decomposition. As shown in Panel B of Table 3, STS displays a sharper and more immediate effect on intraday returns than on overnight returns. The coefficient on STS_{t-1} is negative and highly significant for intraday returns in both markets, indicating that more positive textual sentiment predicts lower returns during the next trading session. The joint tests show that the sum of coefficients from the first through the fifth lag is negative and statistically significant, whereas the sum from the second through the fifth lag is insignificant. No comparable predictability emerges for overnight returns. The absence of overnight predictability suggests that STS-related information is incorporated into prices mainly during trading hours.

Panels A and B also mirror the baseline pattern. Overall, both textual and video-based sentiment are incorporated mainly during trading hours, but STS is associated with a sharper reversal-type response, whereas SVS exhibits a more gradual pattern followed by partial price adjustment over the next several trading days.

[Insert Table 3 about here]

Thirdly, Table 6 examines the joint predictive power of short-video visual sentiment and textual sentiment for market returns. The following specification includes lagged SVS and STS simultaneously, as well as their interaction term, in order to capture whether the two sentiment measures amplify or mitigate each other's pricing effects:

$$R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' STS_{t-i} + \beta'' SVS_{t-1} STS_{t-1} + \gamma X_t + \varepsilon_t, \quad (7)$$

Panel A reports the results for aggregate market returns. When SVS and STS are included jointly, their main effects remain broadly consistent with the baseline results in Table 2 and Table 5. The coefficient on SVS_{t-1} is significantly positive across all three market returns, while the coefficient on SVS_{t-2} is negative and statistically significant for MKT and $SH\ index$. STS continues to display a negative and significant first-lag effect across all specifications. The coefficient on STS_{t-2} is positive and statistically significant only for $SZ\ index$. Overall, the separate dynamics of SVS and STS remain intact: SVS is associated with an initial positive return response followed by partial reversal, whereas STS is associated with an immediate negative response and more limited subsequent correction.

The results also show that the interaction term between SVS_{t-1} and STS_{t-1} is negative and statistically significant. This pattern suggests that the two sentiment measures do not reinforce each other. Instead, they appear to mitigate each other's sentiment-driven pricing effects. To the extent that elevated sentiment tends to push prices away from fundamentals and is subsequently corrected, a stronger reversal can be interpreted as a faster correction of prior overreaction (De Long *et al.* 1990; Obaid & Pukthuanthong 2022). Consistent with this view, when SVS is high, the subsequent reversal associated with STS becomes stronger, suggesting that visual sentiment helps the market correct textual sentiment overreaction more quickly. Conversely, when STS is high, the positive short-run price response associated with SVS is attenuated, indicating that textual sentiment tempers the initial overreaction induced by visual sentiment. In this sense, the two sentiment measures interact by reducing, rather than amplifying, each other's mispricing effects.

The joint significance tests show that the sum of SVS coefficients from the first through the fifth lag remains positive and statistically significant across all three returns, while the sum from the second

through the fifth lag is negative and significant for MKT and $SH\ index$ and weaker for $SZ\ index$. This pattern continues to support a momentum-then-reversal dynamic. For STS, the sum of coefficients from the first through the fifth lag is significantly negative, whereas the sum from the second through the fifth lag is small and insignificant. This indicates that the strong initial negative response associated with STS is not reversed within the next four trading days.

Panel B reports the corresponding results for intraday and overnight returns. As in Table 3, both sentiment measures primarily affect intraday returns rather than overnight returns. For intraday returns, SVS_{t-1} remains positive and significant, whereas STS_{t-1} remains strongly negative and significant in both markets. The interaction term is again significantly negative for both $SH\ intraday$ and $SZ\ intraday$ returns. This result suggests that the offsetting interaction between SVS and STS is concentrated during trading hours, when sentiment is actively incorporated into prices. By contrast, overnight returns show little robust predictability from either sentiment measure, and the interaction effect is generally insignificant.

Taken together, Table 6 suggests that SVS and STS contain distinct but interacting information. Textual sentiment is incorporated into prices more quickly and is associated with a sharper reversal-type response, whereas visual sentiment is incorporated more gradually and is followed by partial price adjustment over the next several trading days. Rather than reinforcing each other, the two forms of sentiment mutually dampen sentiment-driven mispricing: visual sentiment accelerates the correction of textual sentiment-related overreaction, whereas textual sentiment weakens the initial overreaction associated with visual sentiment.

[Insert Table 6 about here]

4. Validation of SVS

To characterize the economic content of SVS, we examine whether it exhibits several empirical patterns commonly associated with investor sentiment. In particular, sentiment should be more state dependent, more relevant for retail-oriented stocks, more sensitive to the trading environment, and more likely to generate disagreement-driven trading when different components of the same signal diverge. We therefore conduct four sets of validation tests along these dimensions.

4.1 Market states

A natural concern is that SVS may capture either sentiment-driven signals or risk-related information. These two interpretations imply different state dependence. If SVS primarily reflects risk, its predictive power should strengthen in down markets, when downside risk is more salient. If, instead, SVS mainly captures investor sentiment, its predictive content should be more pronounced in up markets, when optimistic beliefs are more easily capitalized into prices. To validate the interpretation of SVS, we therefore estimate the following state-dependent specification:

$$R_t = E_t\left(\alpha + \sum_{i=1}^{5}\beta_i SVS_{t-i} + \sum_{i=1}^{5}\beta_i' STS_{t-i} + \beta'' SVS_{t-1}STS_{t-1} + \gamma X_t\right)$$
$$+(1-E_t)\left(A + \sum_{i=1}^{5}B_i SVS_{t-i} + \sum_{i=1}^{5}B_i' STS_{t-i} + B'' SVS_{t-1}STS_{t-1} + \Gamma X_t\right) + \varepsilon_t, \quad (8)$$

where $E_t = 1$ if the one-month-ahead accumulated market return is positive, denoted as an up state, and $E_t = 0$ otherwise, denoted as a down state.

Table 7 shows clear state dependence in the predictive effects of SVS. The coefficient on SVS_{t-1} is positive in both states, but it is larger and more precisely estimated in up markets, except for $SZ\ index$. More importantly, the dynamic response differs sharply across states. In up markets, the sum of coefficients from the second through the fifth lag is significantly negative for MKT and $SH\ index$,

indicating a systematic reversal after the initial positive response. In down markets, by contrast, the subsequent lags do not exhibit a comparable reversal pattern, while the sum of coefficients from the first through the fifth lag remains positive for MKT and $SH\ index$. Thus, the up-state results are characterized mainly by reversal over lags two through five, whereas the down-state results are characterized more by short-horizon persistence over lags one through five. Thus, when market conditions are favorable, visual sentiment generates a stronger initial positive response followed by a more systematic partial reversal. This pattern is more consistent with SVS capturing sentiment-driven signals than risk-related information.

STS shows a different pattern. The coefficient on STS_{t-1} is negative and statistically significant across both up and down states in all three market returns. By contrast, the subsequent STS lags are generally small and imprecisely estimated, and the sum from the second through the fifth lag is insignificant throughout. The cumulative effect of STS is therefore driven mainly by the first lag rather than by a state-dependent sequence of reversals. In this sense, textual sentiment exhibits a more immediate and robust first-lag effect across market states, whereas its state dependence is weaker than that of SVS.

The interaction term between SVS_{t-1} and STS_{t-1} is mostly insignificant across states, with only limited evidence in MKT up states and $SZ\ index$ down states. Thus, the state-dependent results are driven primarily by the main effects of SVS and STS rather than by systematic changes in their interaction across market conditions.

Overall, Table 7 highlights an important asymmetry in sentiment-based predictability. Visual sentiment is more informative in up markets and displays a clearer spike-then-reversal pattern when market condition is favorable, which supports the view that it is more likely to capture sentiment-driven signals than risk-related information. Textual sentiment, by contrast, is characterized by a strong and

immediate first-lag negative effect in both up and down states, suggesting that textual sentiment is incorporated into prices more quickly and in a less state-contingent manner. Taken together, these results indicate that the market response to sentiment depends both on the medium through which sentiment is conveyed and on prevailing market conditions.

[Insert Table 7 about here]

4.2 Retail investor intensity

Given that SVS is constructed from short videos posted by individual influencer accounts, it is naturally more closely related to retail sentiment than to institutionally processed information. If this interpretation is correct, its predictive power should be more visible in stocks with greater retail investor presence. We therefore examine whether the predictive effects of sentiment vary across portfolios with different degrees of retail investor intensity.

Table 8 reports the results. Retail intensity is proxied by stock prices in the cross section: stocks in the top 30%, middle 40%, and bottom 30% of the price distribution are grouped into low-, median-, and high-retail portfolios, respectively, and value-weighted portfolio excess returns are constructed accordingly.

For SVS, the predictive pattern is most clearly visible in the high-retail portfolio. The coefficient on SVS_{t-1} is positive and statistically significant, and the sum of coefficients from the first through the fifth lag remains positive, while the sum from the second through the fifth lag is significantly negative. This pattern indicates an initial positive response followed by partial reversal over subsequent days. A similar but weaker pattern appears in the low-retail portfolio: the first-lag coefficient is also positive and significant, and the subsequent lags imply reversal, but the overall response is less pronounced than in the high-retail portfolio. By contrast, the estimates for the median-retail portfolio are less stable and less

precisely estimated. Importantly, the high-minus-low spread portfolio also exhibits a positive and significant first-lag coefficient on SVS, suggesting that the predictive effect of SVS is stronger in more retail-oriented portfolios. Taken together, these results are consistent with the view that SVS contains sentiment-driven information that is more relevant for stocks with greater retail investor participation.

STS shows a different pattern. The coefficient on STS_{t-1} is negative and statistically significant across all three portfolios, indicating a robust first-lag effect regardless of retail intensity. However, the cumulative predictive content of STS is less uniform across portfolios. The sum of coefficients from the first through the fifth lag is significantly negative in the high- and low-retail portfolios, but not in the median-retail portfolio. Thus, unlike SVS, STS does not display a monotonic relation with retail investor intensity. Instead, its negative first-lag effect is pervasive, while its multi-day predictive effect appears more clearly at the two ends of the retail-intensity distribution.

The interaction term between SVS_{t-1} and STS_{t-1} is negative and statistically significant in the high-minus-low portfolios. This pattern suggests that the offsetting relation between video-based and textual sentiment is stronger where retail participation is highest than lowest. In such settings, each of the two measures appears to dampen the mispricing associated with the other.

Overall, Table 8 highlights meaningful heterogeneity in sentiment-based predictability across retail investor intensity. For SVS, the high-retail portfolio responds more strongly to visual sentiment, exhibiting a clearer momentum-then-partial-reversal pattern, whereas the low-retail portfolio is characterized more by a spike-then-reversal pattern. This contrast suggests that stocks with greater retail investor presence are more sensitive to the information embedded in SVS. By contrast, STS exhibits a less retail-specific predictive pattern: both the high- and low-retail portfolios display persistent reversal dynamics, but without a clear monotonic gradient across retail intensity. These findings reinforce the idea

that SVS and STS contain distinct information, and that retail-oriented stocks are more responsive to the visual component of short-video sentiment.

[Insert Table 8 about here]

4.3 Trading volume

The role of sentiment may also depend on the trading environment. In particular, market-wide trading volume may shape how quickly and how strongly sentiment is reflected in prices. We therefore examine whether the predictive effects of SVS and STS vary across high- and low-volume periods by estimating the following state-dependent specification:

$$R_t = V_t\left(\alpha + \sum_{i=1}^{5}\beta_i SVS_{t-i} + \sum_{i=1}^{5}\beta_i' STS_{t-i} + \beta'' SVS_{t-1}STS_{t-1} + \gamma X_t\right)$$
$$+(1-V_t)\left(\mathrm{A} + \sum_{i=1}^{5}B_i SVS_{t-i} + \sum_{i=1}^{5}B_i' STS_{t-i} + B'' SVS_{t-1}STS_{t-1} + \Gamma X_t\right) + \varepsilon_t, (9)$$

where $V_t = 1$ if lagged market-wide trading volume is above the median, and $V_t = 0$ otherwise.

Table 9 shows that the SVS's first-lag coefficients are positive and statistically significant in both high- and low-volume periods across all market returns. The coefficients are generally larger and more precisely estimated in high-volume states, especially for MKT and $SH\ index$. Subsequent SVS coefficients are mostly negative at short horizons, but the evidence for systematic reversal is limited. The joint tests show that the sum of coefficients from the first through the fifth lag is positive and statistically significant in most cases, whereas the sum from the second through the fifth lag is generally insignificant. Thus, SVS retains a robust short-run predictive effect in both volume states, with stronger and more stable predictability when trading activity is high.

STS exhibits a similarly immediate but negative first-lag effect. The coefficient on STS_{t-1} is negative and statistically significant across both high- and low-volume states in all specifications, but its magnitude is consistently larger in high-volume periods. Unlike the earlier baseline interpretation, the

cumulative predictive content of STS is not stronger in low-volume periods. Instead, the sum of coefficients from the first through the fifth lag is significantly negative in high-volume states, whereas it becomes small and statistically insignificant in low-volume states. The sum from the second through the fifth lag is insignificant throughout. These results indicate that the predictive effect of STS is concentrated in the first lag and is more sustained in high-volume periods.

The interaction term between SVS_{t-1} and STS_{t-1} is generally statistically insignificant in both high- and low-volume periods. This pattern suggests that the volume-state results are driven mainly by the distinct main effects of SVS and STS.

Overall, Table 9 shows that sentiment-based predictability depends not only on the medium through which sentiment is conveyed, but also on the trading environment in which prices are incorporated. Both SVS and STS exhibit stronger and more stable predictive power when market-wide trading volume is high, suggesting that active trading facilitates the incorporation of sentiment into prices. At the same time, the two measures continue to display different dynamic patterns. SVS maintains a positive predictive effect over the one- to five-day horizon in both periods, whereas the negative predictive effect of STS over the same horizon is significant only when trading volume is high. This pattern suggests that visual sentiment contains a more persistent sentiment signal that survives across trading environments, whereas the predictive effect of textual sentiment is more conditional on active trading conditions.

[Insert Table 9 about here]

4.4 Congruence between SVS and STS

We examine whether the congruence between video-based and caption-based sentiment affects trading activity. This test is informative for two reasons. First, if SVS and STS capture sentiment embedded in different components of the same short video, disagreement between them may generate

heterogeneous interpretations and thus affect trading volume. Second, the effect of such disagreement should depend on investor composition. If the divergence between SVS and STS reflects sentiment rather than purely fundamental information, it should be more likely to translate into trading in stocks with greater retail investor presence, where investors are more prone to react to emotionally conveyed signals. To capture both general disagreement and the possibility that large disagreement matters disproportionately, we measure congruence using both $|SVS - STS|$ and $(SVS - STS)^2$ and estimate the following specification:

$$TV_t = \alpha + \theta(SVS - STS)^2{}_{t-1} + \theta'|SVS - STS|_{t-1} + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' STS_{t-i} + \sum_{i=1}^{5} \gamma_i TV_{t-i} + \varepsilon_t, \quad (10)$$

where TV denotes trading volume, constructed as the value-weighted average of stock-level trading volume within each retail-intensity portfolio. This regressions control for lagged trading volume as well as the lagged levels of SVS and STS.

Table 10 reports the results. In the high-retail portfolio, both disagreement measures are positive and highly significant: θ, the coefficient on $(SVS - STS)^2{}_{t-1}$ is 1.090 (t = 5.994), and θ', the coefficient on $|SVS - STS|_{t-1}$ is 1.782 (t = 6.202). These estimates indicate that, in stocks with stronger retail investor presence, greater divergence between video-based and text-based sentiment is associated with higher subsequent trading volume. This pattern is consistent with disagreement-driven trading: when the visual and textual components of the same short video point in different directions, retail investors are more likely to form heterogeneous interpretations and trade on those differences.

The relation weakens and changes sign outside the high-retail portfolio. In the median-retail portfolio, θ remains positive and significant, whereas θ' is negative and significant. This pattern suggests that only sufficiently large divergence stimulates trading, whereas modest divergence may instead reduce

participation. In the low-retail portfolio, both coefficients are negative and statistically significant, indicating that disagreement between SVS and STS is associated with lower subsequent trading volume. This evidence suggests that when retail participation is limited, and institutional participation is correspondingly more important, divergence between video-based and text-based sentiment is less likely to generate active disagreement-driven trading and more likely to be treated as noise or uncertainty.

The coefficients on the lagged levels of SVS and STS reinforce this interpretation. In the high-retail portfolio, both sentiment measures are positively related to subsequent trading volume across several lags, indicating that sentiment-rich short videos stimulate trading activity. This pattern is consistent with the argument that sentiment-induced disagreement leads to higher trading volume (De Long *et al.* 1990; Wang *et al.* 2022). In the low-retail portfolio, by contrast, both SVS and STS are mostly negatively related to turnover, suggesting that when retail participation is weaker, sentiment signals are more likely to be discounted as noise than to trigger active trading demand.

Overall, Table 10 points to a nonlinear and investor-composition-dependent relation between sentiment congruence and trading activity. The fact that disagreement between SVS and STS increases trading mainly in high-retail portfolios suggests that the divergence between the visual and textual components of short-video sentiment captures heterogeneous sentiment-driven interpretations of the same underlying content. By contrast, when retail participation is low, the congruence appears to encourage trading, consistent with conflicting sentiment signals being treated as noise rather than as a source of active disagreement. Taken together, these results provide further support for interpreting SVS and STS as sentiment measures rather than as redundant or purely fundamental signals.

[Insert Table 10 about here]

5. Further analysis

5.1 Institutional vs influencer visual sentiment

As a further comparison, we construct an institutional video sentiment measure using videos posted by institutional accounts and the same extraction procedure as for SVS. This allows us to compare the predictive content of sentiment extracted from institutional videos with that extracted from influencer videos. Specifically, we estimate the following regression:

$$R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' IVS_{t-i} + \beta'' SVS_{t-1} IVS_{t-1} + \gamma X_t + \varepsilon_t, \quad (11)$$

where IVS denotes the video sentiment index.

Table 11 reports the results. Controlling for institutional video sentiment does not alter the main predictive pattern of SVS. The coefficient on SVS_{t-1} remains positive and statistically significant in the aggregate market as well as across the high-, median-, and low-retail portfolios. Moreover, the joint tests continue to support the same dynamic pattern documented earlier: the sum of coefficients on SVS from the first through the fifth lag remains positive and statistically significant, while the sum from the second through the fifth lag is negative and significant in the aggregate market and in the high-retail portfolio. Thus, even after controlling for IVS, influencer video sentiment continues to exhibit a momentum-then-reversal pattern, indicating that its predictive content remains robust.

By contrast, the predictive effect of IVS is much weaker. Although the coefficient on IVS_{t-1} is positive in all specifications, it is only weakly estimated, and the interaction term between SVS_{t-1} and IVS_{t-1} is insignificant throughout. The joint tests likewise show that the cumulative predictive effect of IVS is not statistically significant. These results suggest that institutional video sentiment moves in the same broad direction as influencer video sentiment, but its predictive content is weaker and less stable.

One possible reason is that institutional accounts post substantially fewer videos than individual influencer accounts, which limits the amount of sentiment-rich content entering the market. In our sample, the average number of daily videos is 3.79 for institutional accounts, compared with 19.1 for individual influencer accounts. Another possibility is that institutional videos are more constrained in emotional expression. Relative to influencers, institutional speakers are likely to face stronger reputational, compliance, and communication constraints, making their visual expressions less intense and less variable. As a result, the emotional content embedded in institutional videos may be less salient and harder to extract.

Taken together, Table 11 suggests that sentiment extracted from institutional videos is qualitatively similar to that extracted from influencer videos, but quantitatively weaker in predictive power. This comparison is consistent with the view that sentiment conveyed through the more emotionally expressive and higher-frequency videos posted by individual influencer accounts is more relevant for market pricing, even after controlling for institutional video sentiment.

[Insert Table 11 about here]

5.2 Out-of-sample test

To assess the real-time predictive performance of SVS and STS and to mitigate concerns about in-sample overfitting, we conduct an out-of-sample forecasting exercise following Welch and Goyal (2008). We estimate univariate predictive regressions using only information available up to time *t* to forecast excess returns at *t*+1. The initial estimation window uses the first three months of data, after which the models are re-estimated using a rolling-window scheme as new observations become available. Forecast accuracy is evaluated relative to the historical mean benchmark using the out-of-sample R^2 (R^2_{oos}) statistic of Campbell and Thompson (2008) and the MSFE-adjusted Clark and West (2007) test (CW). In addition, we assess the economic value of the forecasts by computing the annualized Sharpe ratio (SR)

and the certainty equivalent return gain ($CER\ gain$) for a mean–variance investor who allocates wealth between equities and risk-free Treasury bills based on SVS- or STS-based forecasts.

Table 12 reports the out-of-sample forecasting and investment performance of SVS and STS. The results do not indicate uniform dominance of one measure over the other. Instead, SVS and STS appear to contain different types of predictive information, with STS delivering stronger statistical forecast accuracy and SVS generating greater economic value.

For the aggregate market portfolio, STS delivers slightly stronger statistical forecast performance, with a higher R^2_{oos} of 0.409 vs 0.364 for SVS and a larger CW statistic of 4.174 vs 3.188. However, these differences are modest relative to the much larger gap in economic value. The SVS strategy attains an annualized SR of 1.054, about 4.2 times that of STS (0.251), and a $CER\ gain$ of 0.159, about 3.3 times that of STS (0.048). A similar contrast emerges for $SH\ index$. Although SVS and STS perform similarly in statistical forecasting, the SVS-based strategy delivers substantially greater economic value: its Sharpe ratio is more than 12 times that of STS, and its $CER\ gain$ is 0.565, whereas STS generates a negative value of −0.099. Thus, even though textual sentiment performs somewhat better in statistical forecasting, video-based sentiment appears more useful from a practical investment perspective.

For the SZSE market portfolio, STS performs better than SVS in both statistical and economic terms, but the gap in economic value is less pronounced than for the aggregate and SSE market portfolio. STS achieves an R^2_{oos} of 0.227 vs 0.067 for SVS and a CW statistic of 3.876 vs 3.133. Its Sharpe ratio of 0.316 is about 1.5 times that of SVS (0.213), and its $CER\ gain$ of 0.352 is about twice that of SVS (0.176). Thus, while STS has an advantage in the SZSE market, the economic-value gap is still much smaller than the differences in SR and $CER\ gain$ observed for MKT and $SH\ index$.

Overall, the out-of-sample evidence indicates that STS and SVS contain distinct predictive content rather than one uniformly dominating the other. STS tends to perform somewhat better in terms of statistical forecast accuracy in the aggregate and SZSE market, whereas SVS generates much stronger economic value in the aggregate and SSE market. Taken together, these results reinforce the view that visual sentiment and textual sentiment are not redundant: they differ not only in in-sample dynamics, but also in the way their information translates into forecast accuracy and portfolio performance out of sample.

[Insert Table 12 about here]

5.3 Alternative measure of SVS

Following the thin-slices approach of Blankespoor *et al.* (2017), we next focus on the first 30 seconds of each video. The idea is that early segments may capture investors' initial impressions, when attention is likely to be highest and emotionally salient cues may be especially influential. We therefore construct an alternative SVS measure using only the opening 30 seconds of each video (SVS^{30}) and re-estimate the baseline predictive regressions. This test allows us to assess whether sentiment extracted from the most attention-grabbing portion of the video is sufficient to capture the predictive content of the full-video SVS.

Table 13 shows that SVS^{30} yields weaker and less stable predictive patterns than the baseline measure. The individual lag coefficients are generally small and imprecisely estimated across the aggregate market, SSE and SZSE, and their intraday and overnight return components. At the same time, the joint tests on the sum of coefficients from the second through the fifth lag remain significantly negative for MKT, $SH\ index$, $SZ\ index$, and both intraday return series. Thus, the opening segment still captures some reversal-related information, but it fails to reproduce the stronger and more systematic predictive pattern documented for the full-video SVS.

Taken together, Table 13 suggests that sentiment conveyed throughout the full short video contains economically meaningful information beyond that contained in thin slices, consistent with Hu and Ma (2025). While early visual sentiment contributes to return predictability, it does not subsume the predictive content of the baseline SVS. This finding is consistent with the view that sentiment unfolds over the course of a video, and that the market responds not only to initial impressions but also to emotional cues revealed later in the viewing sequence.

[Insert Table 13 about here]

6. Conclusion

This paper shows that short videos on social media convey sentiment that is reflected in equity prices. A facial-expression-based measure of short-video sentiment predicts market returns over short horizons, with an initial positive response followed by partial reversal over the next several trading days. This predictive effect concentrates during trading hours, suggesting that video sentiment is incorporated primarily through intraday trading rather than overnight price setting.

A central takeaway is that video-based sentiment and textual sentiment are not interchangeable. Text-based sentiment exhibits a sharper and more immediate first-lag effect, whereas visual sentiment is incorporated more gradually and displays a more structured spike-then-reversal pattern. When the two measures are entered jointly, their main predictive patterns remain intact, but their interaction terms indicate that they do not simply reinforce each other. Instead, each appears to dampen the mispricing associated with the other, suggesting that investors process the visual and textual components of the same short-video content through distinct channels.

The broader evidence favors an interpretation in which visual sentiment captures sentiment-driven signals rather than fundamentals or risk premia. Its predictive effects are more structured in market up states, more visible in retail-oriented portfolios, and generally stronger in high-volume trading environments. These patterns are difficult to reconcile with a purely risk-based explanation and are more consistent with the view that visual sentiment matters when emotionally conveyed information is more readily transmitted into prices. Additional analyses of visual and textual sentiment congruence and trading volume further show that divergence between the visual and textual components of the same short video has behavioral content: in retail-intensive portfolios, such disagreement is associated with higher subsequent trading volume, consistent with heterogeneous sentiment-driven interpretations, whereas in low-retail portfolios it is more likely to discourage trading.

Our findings have implications for empirical asset pricing and for the study of information transmission through social media in capital markets. As investor attention shifts toward short-form video, measures that directly quantify visual affect can help characterize both the timing and the dynamics of sentiment-driven price responses. More broadly, the evidence suggests that non-textual features of digital media contain market-relevant information beyond what is captured by accompanying text.

Several questions remain open. Future work can examine which aspects of video production—such as editing, vocal tone, visual cues beyond faces, or narrative structure—drive the strongest market responses; how platform mechanisms such as sharing, commenting, and algorithmic recommendation shape diffusion; and whether similar patterns arise in other markets and platforms. Distinguishing investor learning from behavioral trading in response to video sentiment, and extending the analysis to firm-level outcomes in corporate finance and to macro-financial settings, also remain promising directions.

Reference

Andrade, E.B., Odean, T., Lin, S., 2015. Bubbling with Excitement: An Experiment. Review of Finance 20, 447-466

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645-1680

Behrendt, S., Schmidt, A., 2018. The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. Journal of banking & finance 96, 355-367

Blankespoor, E., Hendricks, B.E., Miller, G.S., 2017. Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows. Journal of Accounting Research 55, 275-327

Blankespoor, E., Hendricks, B.E., Miller, G.S., 2023. The Pitch: Managers' Disclosure Choice during Initial Public Offering Roadshows. The Accounting review 98, 1-29

Breaban, A., Noussair, C.N., 2018. Emotional State and Market Behavior. Review of Finance 22, 279-309

Brown, Gregory W., Cliff, Michael T., 2005. Investor Sentiment and Asset Valuation. The Journal of business (Chicago, Ill.) 78, 405-440

Cade, N.L., Koonce, L., Mendoza, K.I., 2020. Using video to disclose forward-looking information: the effect of nonverbal cues on investors' judgments. Review of Accounting Studies 25, 1444-1474

Cai, X., Jiang, F., Kang, J.-K., 2023. Remote Board Meetings and Board Monitoring Effectiveness: Evidence from China. The Review of Financial Studies 36, 4318-4372

Campbell, J.Y., Thompson, S.B., 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? The Review of Financial Studies 21, 1509-1531

Chang, X., Dai, L., Feng, L., Han, J., Shi, J., Zhang, B., 2024. A good sketch is better than a long speech: evaluate delinquency risk through real-time video analysis. Review of Finance 29, 467-500

Chen, J., Tang, G., Yao, J., Zhou, G., 2023. Employee sentiment and stock returns. Journal of Economic Dynamics and Control 149, 104636

Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291-311

Cronin, M., Erkens, D.H., Schloetzer, J.D., Tinsley, C.H., 2021. How Controlling Failure Perceptions Affects Performance: Evidence from a Field Experiment. The Accounting review 96, 205-230

Crowley, R.M., Lou, Y., Tan, S.T., Zhang, L., 2025. Misinformation regulations: early evidence on corporate social media strategy. Review of Accounting Studies 30, 3558-3595

Da, Z., Engelberg, J., Gao, P., 2014. The Sum of All FEARS Investor Sentiment and Asset Prices. The Review of Financial Studies 28, 1-32

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise Trader Risk in Financial Markets. The Journal of political economy 98, 703-738

Edmans, A., Fernandez-Perez, A., Garel, A., Indriawan, I., 2022. Music sentiment and stock returns around the world. Journal of Financial Economics 145, 234-254

Elliott, W.B., Hodge, F.D., Sedor, L.M., 2012. Using Online Video to Announce a Restatement: Influences on Investment Decisions and the Mediating Role of Trust. The Accounting review 87, 513-535

Goldstein, I., Spatt, C.S., Ye, M., 2024. The Next Chapter of Big Data in Finance. The Review of Financial Studies 38, 605-622

Gómez-Cram, R., Grotteria, M., 2022. Real-time price discovery via verbal communication: Method and application to Fedspeak. Journal of Financial Economics 143, 993-1025

Hirshleifer, D., Peng, L., Wang, Q., 2025. News Diffusion in Social Networks and Stock Market Reactions. The Review of financial studies 38, 883-937

Hu, A., Ma, S., 2025. Persuading Investors: A Video-Based Study. The Journal of Finance 80, 2639-2688

Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. Review of Financial Studies 28, 791-837

Huang, S.X., Keskek, S., Sanchez, J.M., 2022. Investor Sentiment and Stock Option Vesting Terms. Management science 68, 773-795

Huang, X., Ivković, Z., Jiang, J.X., Wang, I.Y., 2023. Angel investment and first impressions. Journal of Financial Economics 149, 161-178

Jagolinzer, A.D., 2024. Market and regulatory implications of social identity cohorts: a discussion of crypto influencers. Review of Accounting Studies 29, 2298-2317

Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. Journal of Financial Economics 132, 126-149

Mallipeddi, R.R., Kumar, S., Sriskandarajah, C., Zhu, Y., 2022. A Framework for Analyzing Influencer Marketing in Social Networks: Selection and Scheduling of Influencers. Management science 68, 75-104

Merkley, K.J., Pacelli, J., Piorkowski, M., Williams, B., 2024. Crypto-influencers. Review of Accounting Studies 29, 2254-2297

Nekrasov, A., Teoh, S.H., Wu, S., 2022. Visuals and attention to earnings news on twitter. Review of Accounting Studies 27, 1233-1275

Obaid, K., Pukthuanthong, K., 2022. A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. Journal of Financial Economics 144, 273-297

Pedersen, L.H., 2022. Game on: Social networks and markets. Journal of Financial Economics 146, 1097-1119

Tiwari, A.K., Abakah, E.J.A., Bonsu, C.O., Karikari, N.K., Hammoudeh, S., 2022. The effects of public sentiments and feelings on stock market behavior: Evidence from Australia. Journal of Economic Behavior & Organization 193, 443-472

Wang, W., Su, C., Duxbury, D., 2022. The conditional impact of investor sentiment in global stock markets: A two-channel examination. Journal of Banking & Finance 138, 106458

Welch, I., Goyal, A., 2008. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. The Review of Financial Studies 21, 1455-1508

Wies, S., Bleier, A., Edeling, A., 2023. Finding Goldilocks Influencers: How Follower Count Drives Social Media Engagement. Journal of marketing 87, 383-405

Table 1 Summary statistics

Panel A presents summary statistics, including the observations (N), average ($Mean$), standard deviation ($Std\ dev$), first quartile ($P25$), median ($Median$), and third quartile ($P75$) of short video sentiment (SVS), textual sentiment (STS), the daily returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$), and the corresponding overnight and intraday returns ($SH\ overnight$, $SH\ intraday$, $SZ\ overnight$, $SZ\ intraday$). Panel B reports the Pearson correlation and p-value (p) between SVS and STS. The data covers the trading day spanning from January 2022 to December 2025.

Panel A: Summary statistics of sentiment and market returns

	SVS	STS	MKT	$SH\ index$	$SZ\ index$	$SH\ overnight$	$SH\ intraday$	$SZ\ overnight$	$SZ\ intraday$
N	1212	1212	1212	1212	1212	1212	1212	1212	1212
$Mean$	0.061	0.021	0.078	0.016	0.004	-0.051	0.067	-0.043	0.047
$Std\ dev$	0.039	0.054	1.101	0.989	1.395	0.521	0.896	0.690	1.246
$P25$	0.028	-0.005	-0.479	-0.470	-0.726	-0.222	-0.400	-0.273	-0.662
$Median$	0.055	0.015	0.080	0.020	-0.054	-0.053	0.074	-0.064	0.009
$P75$	0.090	0.048	0.671	0.543	0.736	0.085	0.548	0.160	0.792

Panel B: Correlation between video sentiment and textual sentiment

	STS
SVS	-0.070
p	0.001

Table 2 Short-video visual sentiment and market returns

This table reports the ordinary least squares (OLS) estimates of the β, the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \gamma X_t + \varepsilon_t$. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$). SVS denotes the short-video sentiment index constructed using a large language model (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS is standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)	
	MKT		$SH\ index$		$SZ\ index$	
Variables	β	t	β	t	β	t
SVS_{t-1}	0.215	4.505	0.200	4.617	0.257	4.055
SVS_{t-2}	-0.108	-2.103	-0.106	-2.329	-0.107	-1.606
SVS_{t-3}	-0.076	-1.645	-0.076	-1.857	-0.094	-1.492
SVS_{t-4}	0.035	0.746	0.021	0.516	0.065	1.049
SVS_{t-5}	0.035	0.746	0.041	0.986	0.006	0.100
Sum t-1 to t-5	0.101		0.080		0.127	
Sum t-2 to t-5	-0.114		-0.120		-0.129	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 = 0]	11.678	0.000	9.595	0.001	10.105	0.001
$\chi^2(1)$[Sum t-2 to t-5 = 0]	3.959	0.046	5.465	0.019	3.010	0.082
$Adj.\ R^2$	0.043		0.044		0.032	
N	1207		1207		1207	

Table 3 Short-video sentiment and return decomposition

This table reports the ordinary least squares (OLS) estimates of the β, the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \gamma X_t + \varepsilon_t$ (Panel A), and $R_t = \alpha + \sum_{i=1}^{5} \beta_i STS_{t-i} + \gamma X_t + \varepsilon_t$ (Panel B). R_t denotes overnight and intraday returns in Shanghai Stock Exchange ($SH\ overnight$, $SH\ intraday$), and in Shenzhen Stock Exchange ($SZ\ overnight$, $SZ\ intraday$) on trading day *t*. SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

Panel A: SVS								
	(1)		(2)		(3)		(4)	
	$SH\ overnight$		$SH\ intraday$		$SZ\ overnight$		$SZ\ intraday$	
Variables	β	t	β	t	β	t	β	t
SVS_{t-1}	-0.129	-1.003	0.200	4.985	-0.139	-1.053	0.245	4.202
SVS_{t-2}	0.027	0.942	-0.073	-1.815	0.040	1.105	-0.049	-0.838
SVS_{t-3}	0.210	1.079	-0.087	-2.351	0.222	1.138	-0.109	-1.895
SVS_{t-4}	-0.143	-0.775	0.013	0.351	-0.121	-0.659	0.050	0.899
SVS_{t-5}	-0.100	-1.941	0.023	0.602	-0.126	-2.254	-0.022	-0.401
Sum t-1 to t-5	-0.135		0.076		-0.124		0.115	
Sum t-2 to t-5	-0.006		-0.124		0.015		-0.130	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 = 0]	0.863	0.352	10.76	0.001	0.729	0.393	11.241	0.000
$\chi^2(1)$[Sum t-2 to t-5 = 0]	0.028	0.867	7.509	0.006	0.138	0.710	4.071	0.043
$Adj.\ R^2$	-0.006		0.026		-0.006		0.019	
N	1207		1207		1207		1207	
Panel B: STS								

	(1)		(2)		(3)		(4)	
	$SH\ overnight$		$SH\ intraday$		$SZ\ overnight$		$SZ\ intraday$	
Variable	β	t	β	t	β	t	β	t
STS_{t-1}	-0.013	-0.969	-0.156	-6.309	-0.017	-1.028	-0.231	-6.502
STS_{t-2}	0.093	0.929	0.030	1.270	0.082	0.811	0.063	1.875
STS_{t-3}	0.017	0.591	-0.013	-0.532	0.022	0.688	-0.031	-0.972
STS_{t-4}	-0.062	-0.806	-0.025	-0.956	-0.054	-0.697	-0.044	-1.287
STS_{t-5}	-0.197	-1.044	0.022	0.943	-0.207	-1.087	0.050	1.563
Sum t-1 to t-5	-0.163		-0.142		-0.174		-0.192	
Sum t-2 to t-5	-0.150		0.015		-0.157		0.039	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 = 0]	1.341	0.246	10.525	0.001	1.480	0.223	9.438	0.002
$\chi^2(1)$[Sum t-2 to t-5 = 0]	1.084	0.297	0.128	0.720	1.169	0.279	0.466	0.494
$Adj.\ R^2$	-0.004		0.036		-0.004		0.039	
N	1207		1207		1207		1207	

Table 4 The role of short-video visual sentiment innovation

This table reports the ordinary least squares (OLS) estimates of the β_1, β_1', and β'' (all denoted by β in the table), along with the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' FS_{t-i} + \beta'' SVS_{t-1} FS_{t-1} + \gamma X_t + \varepsilon_t$. R_t denotes market returns on trading day t, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$), and overnight and intraday returns in Shanghai Stock Exchange ($SH\ overnight$, $SH\ intraday$), and in Shenzhen Stock Exchange ($SZ\ overnight$, $SZ\ intraday$). SVS denotes the short-video sentiment index constructed using a large language model (see Section 2.2 for details), whereas FS denotes the volatility of video sentiment innovations measured as the standard deviation of the first difference in second-by-second video sentiment. X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). SVS and FS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
	MKT		$SH\ index$		$SZ\ index$		$SH\ overnight$		$SH\ intraday$		$SZ\ overnight$		$SZ\ intraday$	
Variables	β	t	β	t	β	t	β	t	β	t	β	t	β	t
SVS_{t-1}	0.207	3.771	0.192	3.873	0.244	3.402	-0.154	-0.997	0.194	4.196	-0.162	-1.018	-0.162	-1.018
FS_{t-1}	-0.004	-0.444	-0.009	-0.882	0.014	0.918	0.028	1.745	0.009	0.808	0.034	1.826	0.034	1.826
$SVS_{t-1}FS_{t-1}$	0.011	1.592	0.009	1.307	0.017	1.915	0.018	0.926	0.006	0.827	0.016	0.752	0.016	0.752
Control SVS_{t-2}~SVS_{t-5}	YES		YES		YES		YES		YES		YES		YES	
Control FS_{t-2}~FS_{t-5}	YES		YES		YES		YES		YES		YES		YES	

Adj. R^2	0.041	0.041	0.031	-0.011	0.022	-0.011	-0.011
N	1207	1207	1207	1207	1207	1207	1207

Table 5 Textual sentiment and market returns

This table reports the ordinary least squares (OLS) estimates of the β, the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i STS_{t-i} + \gamma X_t + \varepsilon_t$. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$). STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). STS is standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)	
	MKT		$SH\ index$		$SZ\ index$	
Variables	β	t	β	t	β	t
STS_{t-1}	-0.222	-6.594	-0.187	-6.366	-0.280	-6.362
STS_{t-2}	0.040	1.434	0.030	1.196	0.070	1.970
STS_{t-3}	-0.029	-1.061	-0.018	-0.725	-0.055	-1.610
STS_{t-4}	-0.036	-1.106	-0.034	-1.178	-0.040	-1.003
STS_{t-5}	0.052	1.921	0.046	1.955	0.071	2.009
Sum t-1 to t-5	-0.196		-0.163		-0.234	
Sum t-2 to t-5	0.026		0.024		0.046	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 = 0]	11.086	0.000	10.229	0.001	9.933	0.001
$\chi^2(1)$[Sum t-2 to t-5 = 0]	0.275	0.600	0.300	0.583	0.549	0.458
$Adj.\ R^2$	0.068		0.064		0.061	
N	1207		1207		1207	

Table 6 Joint effects of visual and textual sentiment

This table reports the ordinary least squares (OLS) estimates of the β, β', and β'' (all denoted by β in the table), along with the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' STS_{t-i} + \beta'' SVS_{t-1} STS_{t-1} + \gamma X_t + \varepsilon_t$. In panel A, R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$); and in panel B, R_t denotes overnight and intraday returns in Shanghai Stock Exchange ($SH\ overnight$, $SH\ intraday$), and in Shenzhen Stock Exchange ($SZ\ overnight$, $SZ\ intraday$). SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β, and β' estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

Panel A: Aggregate stock returns

	(1)		(2)		(3)	
	MKT		$SH\ index$		$SZ\ index$	
Variables	β	t	β	t	β	t
SVS_{t-1}	0.192	4.214	0.181	4.356	0.226	3.752
STS_{t-1}	-0.208	-6.200	-0.174	-5.963	-0.264	-6.059

$SVS_{t-1}STS_{t-1}$	-0.041	-2.030	-0.038	-2.074	-0.057	-2.146
SVS_{t-2}	-0.093	-1.875	-0.096	-2.130	-0.086	-1.354
SVS_{t-3}	-0.065	-1.426	-0.066	-1.625	-0.084	-1.371
SVS_{t-4}	0.009	0.188	-0.002	-0.038	0.032	0.528
SVS_{t-5}	0.047	1.025	0.052	1.275	0.025	0.414
STS_{t-2}	0.039	1.416	0.029	1.170	0.069	1.968
STS_{t-3}	-0.027	-1.002	-0.017	-0.676	-0.053	-1.534
STS_{t-4}	-0.036	-1.098	-0.035	-1.203	-0.039	-0.960
STS_{t-5}	0.053	2.006	0.048	2.047	0.071	2.050
Sum t-1 to t-5 SVS	0.090		0.070		0.113	
Sum t-2 to t-5 SVS	-0.102		-0.111		-0.114	
Sum t-1 to t-5 STS	-0.180		-0.150		-0.215	
Sum t-2 to t-5 STS	0.028		0.025		0.049	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	8.735	0.003	7.174	0.007	7.714	0.005
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	3.430	0.064	4.952	0.026	2.459	0.116
$\chi^2(1)$[Sum t-1 to t-5 STS = 0]	9.118	0.002	8.400	0.003	8.201	0.004
$\chi^2(1)$[Sum t-2 to t-5 STS = 0]	0.329	0.566	0.331	0.565	0.624	0.429
Adj. R^2	0.080		0.076		0.072	
N	1207		1207		1207	

Panel B: Intraday and overnight returns

	(1)		(2)		(3)		(4)	
	SH overnight		*SH intraday*		*SZ overnight*		*SZ intraday*	
Variable	β	t	β	t	β	t	β	t
SVS_{t-1}	-0.116	-1.019	0.184	4.722	-0.126	-1.069	0.221	3.950
STS_{t-1}	-0.354	-1.551	-1.439	-5.838	-0.398	-1.581	-2.159	-6.162
$SVS_{t-1}STS_{t-1}$	0.105	0.543	-0.273	-1.690	0.052	0.245	-0.486	-2.013
SVS_{t-2}	0.054	1.093	-0.063	-1.603	0.065	1.203	-0.032	-0.560
SVS_{t-3}	0.200	1.084	-0.078	-2.131	0.214	1.154	-0.100	-1.769

SVS_{t-4}	-0.165	-0.795	-0.005	-0.135	-0.141	-0.684	0.021	0.393
SVS_{t-5}	-0.122	-1.749	0.030	0.793	-0.151	-2.063	-0.007	-0.134
STS_{t-2}	0.858	0.903	0.300	1.281	0.741	0.770	0.633	1.909
STS_{t-3}	0.252	0.703	-0.113	-0.471	0.309	0.812	-0.292	-0.913
STS_{t-4}	-0.687	-0.790	-0.257	-1.011	-0.573	-0.665	-0.421	-1.242
STS_{t-5}	-2.050	-1.060	0.215	0.950	-2.155	-1.105	0.473	1.500
Sum t-1 to t-5 SVS	-0.149		0.068		-0.139		0.103	
Sum t-2 to t-5 SVS	-0.033		-0.116		-0.013		-0.118	
Sum t-1 to t-5 STS	-1.981		-1.294		-2.076		-1.766	
Sum t-2 to t-5 STS	-1.627		0.145		-1.678		0.393	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	0.905	0.341	8.231	0.004	0.780	0.377	9.039	0.002
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	0.346	0.556	7.017	0.008	0.055	0.814	3.542	0.059
$\chi^2(1)$[Sum t-1 to t-5 STS = 0]	1.284	0.257	8.734	0.003	1.380	0.240	8.125	0.004
$\chi^2(1)$[Sum t-2 to t-5 STS = 0]	1.079	0.298	0.13	0.718	1.133	0.287	0.496	0.481
Adj. R^2	-0.004		0.050		-0.004		0.051	
N	1207		1207		1207		1207	

Table 7 The impact of visual sentiment conditional on market states

This table reports the ordinary least squares (OLS) estimates of the β, β', and β'' (all denoted by β in the table), as well as B, B′, B″ (all denoted by B in the table), along with the corresponding t statistics, adjusted R^2s (*Adj.* R^2), and the number of observations (N) for the times series regression: $R_t = E_t\left(\alpha + \sum_{i=1}^{5}\beta_i SVS_{t-i} + \sum_{i=1}^{5}\beta_i' STS_{t-i} + \beta'' SVS_{t-1}STS_{t-1} + \gamma X_t\right) + (1 - E_t)\left(A + \sum_{i=1}^{5}B_i SVS_{t-i} + \sum_{i=1}^{5}B_i' STS_{t-i} + B'' SVS_{t-1}STS_{t-1} + \Gamma X_t\right) + \varepsilon_t$. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$), and we define $E_t = 1$ if the one-month-ahead accumulated market return is positive, denoted as up, and $E_t = 0$ otherwise, denoted as down. SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β, β', B, and B′ estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)				(2)				(3)			
	MKT				$SH\ index$				$SZ\ index$			
	up		down		up		down		up		down	
Variables	β	t	B	t	β	t	B	t	β	t	B	t
SVS_{t-1}	0.231	4.262	0.129	1.537	0.225	4.149	0.129	2.007	0.231	2.842	0.238	2.503

STS_{t-1}	-0.202	-4.260	-0.228	-4.782	-0.186	-3.906	-0.167	-4.766	-0.307	-4.535	-0.249	-4.114
$SVS_{t-1}STS_{t-1}$	-0.047	-1.874	-0.038	-1.055	-0.035	-1.384	-0.036	-1.338	-0.018	-0.511	-0.085	-1.846
SVS_{t-2}	-0.072	-1.269	-0.124	-1.426	-0.074	-1.391	-0.119	-1.675	0.042	0.509	-0.229	-2.295
SVS_{t-3}	-0.105	-1.999	-0.011	-0.135	-0.091	-1.868	-0.042	-0.629	-0.156	-1.836	-0.005	-0.055
SVS_{t-4}	-0.054	-0.988	0.113	1.434	-0.066	-1.308	0.073	1.145	0.004	0.055	0.048	0.490
SVS_{t-5}	0.059	1.044	0.039	0.52	0.057	1.047	0.049	0.817	-0.004	-0.045	0.048	0.561
STS_{t-2}	0.034	0.953	0.042	1.013	0.027	0.714	0.029	0.876	0.111	2.142	0.033	0.684
STS_{t-3}	-0.013	-0.395	-0.046	-1.006	-0.010	-0.323	-0.022	-0.596	-0.037	-0.777	-0.060	-1.248
STS_{t-4}	-0.064	-1.433	0.008	0.158	-0.048	-1.055	-0.017	-0.468	-0.047	-0.944	-0.031	-0.519
STS_{t-5}	0.041	1.216	0.052	1.309	0.032	0.952	0.058	1.864	0.056	1.151	0.089	1.915
Sum t-1 to t-5 SVS	0.059		0.146		0.050		0.090		0.117		0.099	
Sum t-2 to t-5 SVS	-0.172		0.017		-0.175		-0.039		-0.114		-0.139	
Sum t-1 to t-5 STS	-0.203		-0.172		-0.186		-0.119		-0.225		-0.218	
Sum t-2 to t-5 STS	-0.002		0.057		0.000		0.048		0.083		0.031	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	2.624	0.105	7.218	0.007	2.038	0.153	4.995	0.025	6.294	0.012	2.361	0.124
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	5.874	0.015	0.032	0.857	6.386	0.011	0.315	0.574	1.506	0.219	1.484	0.223
$\chi^2(1)$[Sum t-1 to t-5 STS = 0]	6.317	0.011	3.409	0.064	5.258	0.021	3.223	0.072	5.708	0.016	3.641	0.056
$\chi^2(1)$[Sum t-2 to t-5 STS = 0]	0.001	0.979	0.480	0.488	0.000	0.996	0.661	0.416	1.113	0.291	0.121	0.727
Adj. R^2	0.076				0.071				0.068			
N	1185				1185				1185			

Table 8 The impact of visual sentiment across retail investor intensity

This table reports the ordinary least squares (OLS) estimates of the β, β', and β'' (all denoted by β in the table), along with the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' STS_{t-i} + \beta'' SVS_{t-1} STS_{t-1} + \gamma X_t + \varepsilon_t$. R_t denotes the daily excess return of value-weighted portfolios formed by sorting stocks according to the cross-sectional distribution of stock prices. Stocks in the bottom 30%, middle 40%, and top 30% of prices are grouped into high-, median-, and low-retail portfolios, respectively. SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β, and β' estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)		(4)	
	high		median		low		high-low	
Variables	β	t	β	t	β	t	β	t
SVS_{t-1}	0.236	4.357	0.447	1.584	0.141	3.312	0.096	2.150
STS_{t-1}	-0.236	-6.064	-0.181	-3.073	-0.153	-5.158	-0.087	-3.075
$SVS_{t-1}STS_{t-1}$	-0.056	-2.241	-0.067	-1.652	-0.018	-0.943	-0.038	-1.683
SVS_{t-2}	-0.090	-1.534	-0.180	-1.810	-0.100	-2.221	0.010	0.215
SVS_{t-3}	-0.076	-1.381	-0.443	-1.142	-0.070	-1.806	-0.024	-0.494

SVS_{t-4}	0.030	0.549	0.393	0.976	-0.019	-0.476	0.050	1.069
SVS_{t-5}	0.016	0.296	0.185	1.427	0.089	2.345	-0.048	-1.064
STS_{t-2}	0.062	1.855	-0.161	-0.809	0.001	0.045	0.046	1.558
STS_{t-3}	-0.049	-1.541	-0.107	-1.186	0.007	0.281	-0.043	-1.598
STS_{t-4}	-0.044	-1.180	0.161	0.876	-0.020	-0.659	-0.018	-0.665
STS_{t-5}	0.066	2.071	0.486	1.109	0.034	1.425	0.020	0.685
Sum t-1 to t-5 SVS	0.116		0.402		0.041		0.084	
Sum t-2 to t-5 SVS	-0.120		-0.045		-0.100		-0.012	
Sum t-1 to t-5 STS	-0.201		0.198		-0.131		-0.082	
Sum t-2 to t-5 STS	0.035		0.379		0.022		0.005	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	9.608	0.001	1.539	0.214	2.548	0.110	8.084	0.004
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	3.361	0.066	0.232	0.629	4.050	0.044	0.048	0.825
$\chi^2(1)$[Sum t-1 to t-5 STS = 0]	8.554	0.003	0.262	0.608	6.123	0.013	2.475	0.115
$\chi^2(1)$[Sum t-2 to t-5 STS = 0]	0.389	0.533	1.195	0.274	0.230	0.631	0.010	0.918
Adj. R^2	0.076		-0.001		0.060		0.019	
N	1207		1207		1207		1207	

Table 9 The impact of visual sentiment conditional on market-wide trading volume

This table reports the ordinary least squares (OLS) estimates of the β, β', and β'' (all denoted by β in the table), as well as B, B′, B″ (all denoted by B in the table), along with the corresponding t statistics, adjusted R^2s (*Adj.* R^2), and the number of observations (N) for the times series regression: $R_t = V_t\left(\alpha + \sum_{i=1}^{5}\beta_i SVS_{t-i} + \sum_{i=1}^{5}\beta_i' STS_{t-i} + \beta'' SVS_{t-1}STS_{t-1} + \gamma X_t\right) + (1 - V_t)\left(A + \sum_{i=1}^{5} B_i SVS_{t-i} + \sum_{i=1}^{5} B_i' STS_{t-i} + B'' SVS_{t-1}STS_{t-1} + \Gamma X_t\right) + \varepsilon_t$. We define $V_t = 1$ for months with lagged market-wide trading volume above the median, denoted as high, and $V_t = 0$ otherwise, denoted as low. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$). SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β, β', B, and B′ estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)				(2)				(3)			
	MKT				$SH\ index$				$SZ\ index$			
	high		low		high		low		high		low	
Variables	β	t	B	t	β	t	B	t	β	t	B	t
SVS_{t-1}	0.246	3.123	0.131	2.478	0.249	3.455	0.110	2.301	0.290	2.760	0.158	2.229

STS_{t-1}	-0.290	-4.583	-0.111	-3.572	-0.233	-4.206	-0.103	-3.785	-0.388	-4.707	-0.123	-3.099
$SVS_{t-1}STS_{t-1}$	-0.047	-1.216	0.016	0.643	-0.048	-1.408	0.009	0.419	-0.051	-1.003	0.018	0.569
SVS_{t-2}	-0.156	-1.744	-0.057	-1.093	-0.169	-2.068	-0.045	-0.957	-0.114	-1.002	-0.079	-1.184
SVS_{t-3}	-0.143	-1.367	-0.033	-0.752	-0.141	-1.536	-0.036	-0.911	-0.190	-1.360	-0.035	-0.593
SVS_{t-4}	-0.016	-0.176	0.045	0.919	-0.022	-0.286	0.031	0.701	-0.022	-0.181	0.088	1.415
SVS_{t-5}	0.179	1.872	-0.010	-0.225	0.179	2.094	-0.005	-0.125	0.181	1.482	-0.042	-0.711
STS_{t-2}	0.013	0.289	0.047	1.492	-0.002	-0.055	0.043	1.549	0.068	1.165	0.056	1.338
STS_{t-3}	-0.033	-0.732	-0.022	-0.696	-0.026	-0.636	-0.008	-0.283	-0.079	-1.373	-0.031	-0.805
STS_{t-4}	-0.043	-0.728	-0.027	-0.942	-0.048	-0.900	-0.022	-0.856	-0.038	-0.516	-0.038	-0.987
STS_{t-5}	0.046	1.016	0.050	1.652	0.031	0.796	0.052	1.926	0.059	1.003	0.069	1.770
Sum t-1 to t-5 SVS	0.110		0.076		0.095		0.055		0.146		0.090	
Sum t-2 to t-5 SVS	-0.136		-0.055		-0.154		-0.056		-0.144		-0.068	
Sum t-1 to t-5 STS	-0.308		-0.064		-0.278		-0.039		-0.377		-0.067	
Sum t-2 to t-5 STS	-0.018		0.047		-0.045		0.065		0.011		0.055	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	4.781	0.028	4.600	0.031	4.462	0.034	3.025	0.081	4.480	0.034	4.029	0.044
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	1.960	0.161	0.884	0.347	3.178	0.074	1.059	0.303	1.273	0.259	0.759	0.383
$\chi^2(1)$[Sum t-1 to t-5 STS = 0]	6.815	0.009	1.478	0.224	7.460	0.006	0.713	0.398	6.571	0.010	0.984	0.321
$\chi^2(1)$[Sum t-2 to t-5 STS = 0]	0.036	0.850	0.796	0.372	0.309	0.577	1.967	0.160	0.009	0.924	0.652	0.419
Adj. R^2	0.083				0.080				0.075			
N	1207				1207				1207			

Table 10 Congruence between visual and textual sentiment and trading volume

This table reports the ordinary least squares (OLS) estimates of the β, β', θ, and θ' (all denoted by β in the table), along with the corresponding t statistics, adjusted R^2s (*Adj.* R^2), and the number of observations (N) for the times series regression: $TV_t = \alpha + \theta(SVS - STS)^2{}_{t-1} + \theta'|SVS - STS|_{t-1} + \sum_{i=1}^{5}\beta_i SVS_{t-i} + \sum_{i=1}^{5}\beta_i' STS_{t-i} + \sum_{i=1}^{5}\gamma_i TV_{t-i} + \varepsilon_t$. Columns (1)–(3) report results for portfolios across different retail intensities. Portfolios are constructed based on the cross-sectional distribution of stock prices, with stocks in the bottom 30%, middle 40%, and top 30% of prices grouped into low-, median-, and high-retail portfolios, and portfolio trading volume (TV_t) is computed on a value-weighted basis. SVS denotes the short-video sentiment index constructed using a large language model, whereas STS denotes the sentiment extracted from short-video texts using a machine-learning method (see Section 2.2 for details). Accordingly, their absolute difference $|SVS - STS|_{t-1}$ and square of difference $(SVS - STS)^2{}_{t-1}$ measure the congruence between SVS and STS. Five-day lagged trading volume (TV_{t-1}~TV_{t-5}) are controlled in this regression. X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). SVS and STS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)	
	high		median		low	
Variables	β	t	β	t	β	t
$(SVS - STS)^2{}_{t-1}$	1.090	5.994	0.122	2.512	-0.484	-2.637
$\|SVS - STS\|_{t-1}$	1.782	6.202	-0.240	-2.567	-1.323	-3.910
SVS_{t-1}	0.636	3.591	-0.123	-2.324	-0.871	-4.654
SVS_{t-2}	0.778	4.192	0.009	0.180	-0.654	-3.977
SVS_{t-3}	0.846	5.012	-0.014	-0.280	-0.716	-3.981
SVS_{t-4}	0.993	5.183	-0.011	-0.212	-0.998	-5.557

SVS_{t-5}	-0.204	-2.459	0.031	1.276	0.136	1.290
STS_{t-1}	0.352	2.830	-0.021	-0.622	-0.319	-2.775
STS_{t-2}	0.352	3.045	-0.055	-1.673	-0.324	-2.663
STS_{t-3}	0.336	3.320	-0.020	-0.642	-0.204	-1.717
STS_{t-4}	0.221	2.171	-0.057	-1.819	-0.293	-2.702
STS_{t-5}	0.262	2.634	-0.062	-1.939	-0.393	-3.785
Control TV_{t-1}~TV_{t-5}	YES		YES		YES	
Adj. R^2	0.643		0.569		0.491	
N	1207		1207		1207	

Table 11 Individual and institutional visual sentiment

This table reports the ordinary least squares (OLS) estimates of the β_1, β_1', and β'' (all denoted by β in the table), along with the corresponding t statistics, adjusted R^2s (*Adj.* R^2), and the number of observations (N) for the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i} + \sum_{i=1}^{5} \beta_i' IVS_{t-i} + \beta'' SVS_{t-1} IVS_{t-1} + \gamma X_t + \varepsilon_t$. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT, column (1)), and daily excess return of value-weighted portfolios formed by sorting stocks according to the cross-sectional distribution of stock prices (column (2)—(4)). Stocks in the bottom 30%, middle 40%, and top 30% of prices are grouped into high-, median-, and low-retail portfolios, respectively. SVS denotes the short-video sentiment index constructed using a large language model (see Section 2.2 for details), whereas IVS denotes the video sentiment index constructed from institutional videos using the same method. X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β, and β' estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS and IVS are standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)		(4)	
	MKT		high		median		low	
Variables	β	t	β	t	β	t	β	t
SVS_{t-1}	0.226	4.625	0.242	4.825	0.209	3.989	0.202	3.640
IVS_{t-1}	0.907	1.662	0.714	1.157	1.067	1.820	0.708	1.338
$SVS_{t-1}IVS_{t-1}$	-0.072	-0.207	0.178	0.516	-0.026	-0.071	-0.366	-0.872

Control SVS_{t-2}~SVS_{t-5}	YES		YES		YES		YES	
Control IVS_{t-2}~IVS_{t-5}	YES		YES		YES		YES	
Sum t-1 to t-5 SVS	0.117		0.089		0.110		0.111	
Sum t-2 to t-5 SVS	-0.109		-0.153		-0.099		-0.091	
Sum t-1 to t-5 IVS	0.604		0.973		0.813		0.029	
Sum t-2 to t-5 IVS	-0.303		0.259		-0.254		-0.679	
	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 SVS = 0]	9.658	0.001	5.369	0.020	7.022	0.008	6.602	0.010
$\chi^2(1)$[Sum t-2 to t-5 SVS = 0]	3.366	0.066	6.276	0.012	2.431	0.118	1.944	0.163
$\chi^2(1)$[Sum t-1 to t-5 IVS = 0]	0.610	0.434	1.335	0.247	0.895	0.344	0.001	0.972
$\chi^2(1)$[Sum t-2 to t-5 IVS = 0]	0.149	0.699	0.094	0.759	0.088	0.766	0.633	0.426
Adj. R^2	0.041		0.047		0.033		0.033	
N	1207		1207		1207		1207	

Table 12 Out-of-sample performance

This table reports the out-of-sample R^2 (R^2_{oos}), Clark and West (2007) test statistic (CW), annualized Sharpe ratio (SR), and certainty equivalent return gain ($CER\ gain$) for a mean–variance investor with a relative risk aversion of 2. The investor allocates wealth monthly between equities and risk-free Treasury bills based on forecasts from univariate predictive regressions. These regressions use lagged SVS or STS to predict excess returns for the value-weighted stock market portfolio (MKT in panel A), the Shanghai Stock Exchange Composite Index ($SH\ index$ in panel B), and the Shenzhen Stock Exchange Component Index ($SZ\ index$ in panel C). The CER gain is the return difference between using the predictor-based forecast and the prevailing mean forecast. The equity weight is constrained to lie between −1 and 2. SVS and STS are standardized to have a zero mean and unit variance. The out-of-sample period spans the trading days from January 2022 to December 2022.

Panel A: MKT				
	R^2_{oos}	CW	SR	$CER\ gain$
SVS	0.364	3.188	1.054	0.159
STS	0.409	4.174	0.251	0.048
Panel B: $SH\ index$				
	R^2_{oos}	CW	SR	$CER\ gain$
SVS	0.442	3.010	0.602	0.565
STS	0.418	3.863	0.048	-0.099
Panel C: $SZ\ index$				
	R^2_{oos}	CW	SR	$CER\ gain$
SVS	0.067	3.133	0.213	0.176
STS	0.227	3.876	0.316	0.352

Table 13 Alternative visual sentiment measure within the head 30 seconds

This table reports the results with an alternative visual sentiment index constructed using the head 30 seconds of the short videos (SVS^{30}). The ordinary least squares (OLS) estimates of the β, the corresponding t statistics, adjusted R^2s ($Adj.\ R^2$), and the number of observations (N) are obtained from the times series regression: $R_t = \alpha + \sum_{i=1}^{5} \beta_i SVS_{t-i}^{30} + \gamma X_t + \varepsilon_t$. R_t denotes market returns on trading day *t*, including the daily excess returns on the value-weighted stock market portfolio (MKT), the Shanghai Stock Exchange Composite Index ($SH\ index$), and the Shenzhen Stock Exchange Component Index ($SZ\ index$), and overnight and intraday returns in Shanghai Stock Exchange ($SH\ overnight,\ SH\ intraday$), and in Shenzhen Stock Exchange ($SZ\ overnight,\ SZ\ intraday$). X is a spectrum of control variables, including the five lags of both market returns and square of market returns, and day-of-the-week dummies (except for Monday). This table also reports the sum of β estimates, and the corresponding chi-square statistics ($\chi^2(1)$) and p-value (p). SVS is standardized to have a zero mean and unit variance. The data covers the trading day spanning from January 2022 to December 2025.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
	MKT		$SH\ index$		$SZ\ index$		$SH\ overnight$		$SH\ intraday$		$SZ\ overnight$		$SZ\ intraday$	
Variables	β	t	β	t	β	t	β	t	β	t	β	t	β	t
SVS_{t-1}^{30}	-0.046	-0.646	-0.037	-0.571	-0.065	-0.715	0.142	0.993	-0.043	-0.684	0.144	0.987	-0.061	-0.713
SVS_{t-2}^{30}	0.014	0.407	0.019	0.610	-0.001	-0.021	-0.033	-0.808	0.013	0.477	-0.037	-0.811	-0.024	-0.636
SVS_{t-3}^{30}	-0.004	-0.158	-0.006	-0.241	0.015	0.429	-0.072	-1.007	-0.003	-0.109	-0.078	-1.054	0.026	0.731
SVS_{t-4}^{30}	-0.023	-0.785	-0.014	-0.573	-0.038	-1.001	0.115	0.845	-0.009	-0.403	0.105	0.777	-0.029	-0.826
SVS_{t-5}^{30}	-0.048	-1.825	-0.046	-1.986	-0.048	-1.358	0.004	0.194	-0.038	-1.666	0.013	0.607	-0.038	-1.100
Sum t-1 to t-5	-0.107		-0.085		-0.138		0.156		-0.080		0.147		-0.126	
Sum t-2 to t-5	-0.061		-0.048		-0.073		0.014		-0.037		0.003		-0.065	

	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
$\chi^2(1)$[Sum t-1 to t-5 = 0]	9.405	0.398	7.752	0.416	8.675	0.486	0.842	0.964	8.065	0.294	0.727	0.954	9.656	0.286
$\chi^2(1)$[Sum t-2 to t-5 = 0]	0.725	0.011	0.541	0.033	0.642	0.009	0.051	0.320	0.363	0.007	0.002	0.369	0.620	0.003
Adj. R^2	0.029		0.029		0.021		-0.006		0.009		-0.007		0.007	
N	1207		1207		1207		1207		1207		1207		1207	