Visual Presentation, Investor Attention, and Market Reactions to Financial YouTube Content

Abstract

This paper examines how financial YouTube content affects investor attention and stock market reactions. We construct a novel dataset linking videos from 25 major finance-focused YouTube channels during 2021–2024 to firm-level market outcomes using computer vision, natural language processing, and audio transcription. An event-study analysis shows that visually salient thumbnails and emotionally charged titles are associated with higher trading activity, increased volatility, and short-run return movements that often partially reverse. Transcript sentiment exhibits weaker and less consistent effects, while influencer reputation amplifies investor participation and disagreement. Overall, the evidence suggests that visual and linguistic features of financial videos influence market outcomes through attention and belief-dispersion channels, highlighting the growing role of multimodal digital media in financial markets.

Keywords: YouTube, Social Media, Investor Attention, Visual Design, Textual Framing, Information Diffusion, Belief Dispersion, Trading Activity

JEL Classification : G14, G41, G12, D83, L82

1 Introduction

Retail investors increasingly encounter firm-specific financial information through digital platforms rather than through traditional disclosures, analyst reports, or business news. YouTube is a central part of this information environment. Finance-focused creators package market commentary, firm analysis, earnings reactions, and investment narratives in videos that are designed not only to inform viewers but also to attract attention. Unlike written news or corporate filings, however, YouTube content is encountered first through visual and stylistic cues: thumbnails, facial expressions, colors, typography, and titles. These features shape whether investors click on a video and may also shape how they interpret its message.

This paper studies whether the visual and linguistic features of financial YouTube videos are associated with stock market reactions. The question is motivated by a large literature showing that investor attention affects trading and prices. Retail investors are more likely to buy attention-grabbing stocks (Barber and Odean, 2008), search-based attention predicts trading activity and short-lived price pressure (Da et al., 2011), and limited attention affects how investors process financial information (Hirshleifer and Teoh, 2003; Hirshleifer et al., 2009). Related work shows that media tone and textual sentiment contain information about market reactions (Tetlock, 2007; Tetlock et al., 2008; Loughran and McDonald, 2016). Yet most empirical finance research measures attention and sentiment using text, search activity, or news coverage. Much less is known about whether visual presentation affects investor behavior. This distinction is especially relevant because social-media platforms do not merely transmit information; they also shape how attention is allocated, diffused, and amplified across investors (Rakowski et al., 2021; Paydarzarnaghi and Rakowski, 2026).

The omission matters because video platforms are multimodal. A viewer often sees the thumbnail and title before hearing any spoken analysis. These pre-consumption cues may be especially relevant when investors face cognitive constraints and rely on heuristics. Visual stimuli can activate rapid intuitive processing (Kahneman, 2011), affective judgments (Slovic et al., 2002), and emotion-based attention (Lang, 2000). In a financial setting, such cues may attract retail attention, induce disagreement, and generate short-run trading pressure even when the underlying verbal information is limited. At the same time, transcript sentiment may convey more substantive information, while linguistic complexity may affect processing costs and belief dispersion (Bloomfield, 2002; Li, 2008; Miller, 2010).

We construct a novel event-level dataset linking finance-related YouTube videos to firm-level market outcomes. Each video is linked to firms using ticker detection, firm-name matching, and semantic matching. We extract visual features from thumbnails, textual features from titles and optical-character-recognition output, and spoken content from transcripts generated using Whisper (Radford et al., 2023). We then align each YouTube–firm event to the first trading day on which investors could respond to the video. The final analysis sample contains 1,924 retained videos posted between 2021 and 2024, which yield 12,492 video–firm event observations across 87 publicly traded firms.

Our main explanatory variables capture four dimensions of video content. First, thumbnail sentiment measures visual salience and affective intensity, including facial expressions, color features, and image-based text. Second, title sentiment captures emotional and promotional framing in the video title. Third, transcript sentiment measures the tone of the spoken narrative using both general and finance-specific sentiment dictionaries. Fourth, textual richness captures linguistic complexity and information density. We also measure influencer reputation using creator-level characteristics such as audience size and engagement. These variables allow us to separate visual attention, textual information, cognitive complexity, and creator reach within the same empirical setting.

The empirical design is an event-time panel framework. We estimate market reactions around video release dates using abnormal returns, abnormal trading volume, and abnormal volatility over short and medium horizons. Because some videos are posted during regular trading hours, same-day outcomes may partly reflect trading before the video was released. We therefore treat all $t = 0$ estimates as contemporaneous benchmarks against which we can evaluate the extent of subsequent reversals, following Barber and Odean (2008). Our main tests and interpretations focus on outcomes beginning at $t + 1$, including cumulative windows over the first trading week, month, and quarter after the video release. The design does not require the assumption that video posting is exogenous. Instead, it asks whether, conditional on firm fixed effects, time fixed effects, firm characteristics, and other content attributes, the visual and linguistic features of videos are associated with differential market reactions after release. We also examine placebo windows before video releases, alternative event windows, and matched-sample specifications to assess whether the results are driven by pre-existing trends or observable firm characteristics.

The evidence points to three main findings. First, visual and title-based cues are associated primarily with investor attention. More visually salient or emotionally framed videos are followed by (i.e., from t+1 and onward) higher trading activity and, in several specifications, higher volatility. Return effects are short-lived and often partially reverse, consistent with temporary attention-driven price pressure rather than persistent changes in firm value. Second, transcript sentiment has weaker and less stable effects than visual presentation. This pattern suggests that spoken sentiment is not the only channel through which video content affects markets. Third, influencer reputation amplifies trading participation and is associated with greater disagreement in some specifications, indicating that creator reach affects how widely content is processed by investors.

The paper contributes to three literatures. First, it extends the investor-attention literature by showing that attention can be generated by visual design, not only by news coverage, search activity, or textual salience. Second, it contributes to research on media sentiment and asset prices by moving beyond text-based measures toward a multimodal view of financial communication. Third, it adds to work on social media and retail trading by studying YouTube creators as informal financial intermediaries whose presentation choices may affect attention, disagreement, and short-run price formation.

The paper also has implications for market efficiency. The results suggest that digital financial media can affect trading and volatility through packaging and presentation, even when return effects are temporary. This does not imply that YouTube videos contain no information or that visual content mechanically causes mispricing. Rather, the evidence indicates that modern financial communication combines information with attention engineering. Understanding how investors respond to that combination is increasingly relevant as firm-specific information is filtered through creator-driven and algorithmically curated media environments.

2 Hypothesis Development

Digital platforms such as YouTube increasingly function as major information and attention channels for retail investors. Unlike traditional disclosures, YouTube videos embed information across multiple modalities—visual design, textual framing, and spoken narrative—each of which activates different cognitive, emotional, and behavioral mechanisms. A large behavioral-finance literature shows that investors allocate attention imperfectly and rely on heuristics when processing information. Attention constraints (Hirshleifer et al., 2009), salience-driven belief distortions (Bordalo et al., 2020), and affective judgments (Slovic et al., 2002) generate short-term market reactions even in the absence of new fundamentals.

YouTube provides an environment in which these mechanisms become amplified. Thumbnails and titles act as the *initial point of contact* between content creators and viewers. These elements are intentionally engineered to be visually striking—featuring bright colors, exaggerated facial expressions, and emotionally charged wording—in order to attract clicks and increase view-through rates, which in turn affects platform-level promotion dynamics (Koh and Cui, 2022). Because visual cues are processed rapidly through intuitive (System 1) channels (Kahneman, 2011), they may influence investor attention before any substantive information is absorbed.

By contrast, the transcript conveys substantive information about firms, events, or market conditions. A rich literature shows that textual sentiment and language tone can reveal private signals, shape investor beliefs, and affect asset prices (Tetlock, 2007; Tetlock et al., 2008; Hassan et al., 2019). These effects typically unfold gradually as investors interpret the content, leading to return continuation or reductions in uncertainty.

A third modality arises from linguistic complexity. Dense, technical, or information-rich language increases cognitive load and widens belief dispersion among heterogeneous investors (Bloomfield, 2002; Li, 2008). Such frictions can generate higher volatility and disagreement-based trading.

Finally, investor responses may depend on the identity of the creator. Creators with large followings, established reputations, or perceived expertise enjoy stronger social credibility. Research on media influence and attention cascades (Ahern and Sosyura, 2014; Hirshleifer et al., 2025) suggests that trusted or widely followed communicators can amplify market reactions by reaching larger audiences and shaping interpretation.

Taken together, these observations motivate a multimodal framework in which *visual attention*, *textual information*, *linguistic complexity*, and *social amplification* jointly shape how markets respond to financial YouTube content. These considerations lead to the following testable hypotheses.

H1 (Visual Attention Hypothesis). Videos with stronger thumbnail and title visual intensity generate more immediate market reactions through attention-driven trading.

H2 (Textual Information Hypothesis). More positive transcript sentiment conveys fundamental information that markets incorporate gradually, tending to produce return continuation and, in some cases, reduced uncertainty.

H3 (Linguistic Complexity Hypothesis). Greater textual richness increases cognitive load and belief dispersion, leading to delayed price adjustment and potentially greater heterogeneity in market responses.

H4 (Influencer Amplification Hypothesis). Videos from more reputable or widely followed creators influence the magnitude and nature of investor responses, reflecting broader reach and differences in perceived credibility.

3 Data

We collect videos from Jan 2021–Dec 2024 for 25 top finance-focused YouTubers. We identified the 25 finance-focused YouTube channels with the largest subscriber bases using Social Blade. Subscriber count was used as an objective measure of channel influence and audience reach.

Each video is mapped to tickers via explicit mentions, company-name matches, or embedding-based matching. Metadata includes views, likes, and comments. We align each video to the trading day on which the information first becomes tradeable by the market. Videos posted during regular trading hours (9:30–16:00 Eastern Time) are assigned to the same-day trading date as the event day $t = 0$. Videos posted after market close or on non-trading days (weekends and market holidays) are assigned to the next CRSP trading day. This timing convention follows standard practice in the event-study literature (Barber and Odean, 2008; Da et al., 2011; Fang and Peress, 2009).

3.1 YouTube Video Collection, Filtering, and Transcript Construction

The empirical analysis relies on a multi-year panel of financial YouTube content posted between January 2021 and December 2024. Retail investors increasingly consume video-based financial commentary, and prior work shows that attention to online media can meaningfully influence trading behavior, liquidity, and price reactions (Barber and Odean, 2008; Da et al., 2011; Fang and Peress, 2009). To construct the dataset, we identify twenty-five prominent creators who specialize in equity analysis, valuation discussions, earnings breakdowns, and market commentary. These channels collectively represent some of the largest and most influential hubs in financial YouTube. Table A-6 lists the 25 channels in the sample, their approximate subscriber counts, and their principal content focus. The sample is intended to capture prominent English-language channels that regularly discuss individual publicly traded equities; it is not intended to represent the universe of finance-related YouTube content.

Using the YouTube Data API, we retrieve the full historical archive for each channel, including titles, publication timestamps, engagement statistics, and thumbnail images. Each observation is uniquely indexed by the tuple {channel_name, video_event_id, published_at}. To ensure that the analyzed content is meaningful for firm-level event studies, the sample was restricted to stock market–related videos. Videos primarily focused on cryptocurrencies, live streams, YouTube shorts, premieres, story-format uploads, market recaps, or other non-stock-related content were excluded, as were videos with insufficient spoken content. These filtering steps remove formats unlikely to convey economically relevant information. Videos are required to mention at least two distinct stocks, ensuring that transcript-based sentiment is measured for multi-firm content rather than single-ticker-only videos. Table A-1 reports the progression from the raw video archive to the final event-level analysis sample. We begin with 3,609 raw videos and retain 1,924 videos after applying the content and quality screens. These videos generate 14,745 preliminary video–firm links. After merging with CRSP and Compustat and imposing the data-availability requirements for the event-study analysis, the final sample contains 12,492 video–firm events across 87 firms.

All retained videos are transcribed using Whisper, a state-of-the-art neural speech- recognition model that is robust to background noise, varied speaking styles, and spontaneous phrasing (Radford et al., 2023). The transcription output includes segment-level timestamps and a consolidated full-text transcript suitable for natural-language processing. This uniform treatment ensures comparability across creators with different production styles.

To link videos to publicly traded firms, we employ a hybrid mapping strategy combining (i) explicit ticker detection in titles and transcripts, (ii) canonical firm-name recognition, and (iii) embedding-based semantic similarity derived from transformer models (Vaswani et al., 2017). Embedding-based matching captures colloquial references (e.g., “the fruit company” for Apple), reducing false negatives common in text-only identification systems. The outcome is a set of YouTube–Firm Events, each representing a unique instance in which a firm is discussed.

All continuous YouTube-derived features (sentiment, visual intensity, linguistic features, engagement) are winsorized at the 1st and 99th percentiles to mitigate the impact of likely transcription or classification errors. Log-transformed variables are not winsorized, as the transformation naturally reduces the impact of extreme values.

3.2 Thumbnail, Title, and Visual Feature Engineering

Thumbnails and titles constitute the primary entry points for capturing viewer attention, consistent with evidence that visual and affective salience strongly influences investor behavior (Bordalo et al., 2020; Slovic et al., 2002). Each thumbnail is extracted and processed using computer-vision algorithms that quantify saturation, hue distribution, luminosity, contrast gradients, and pixel-level salience. These characteristics capture the creator's intended visual design choices.

Human faces in thumbnails are detected using established face-recognition models, and each face is evaluated along emotional dimensions such as fear, excitement, anger, surprise, and exaggerated expressive cues used in attention-maximizing content. The resulting emotion vectors yield a Visual Sentiment Score capturing affective arousal.

Thumbnail text is extracted using OCR, and both OCR text and video titles are processed for hyperbolic cues such as ALL-CAPS usage, exclamation marks, and extreme sentiment words (e.g., "CRASH," "URGENT," "INSANE"). These linguistic markers proxy for deliberate attention engineering. The combination of these image-based and text-based features provides a multidimensional representation of visual salience.

3.3 Validation of Visual Measures

We validate the visual measures in two ways. First, we examine whether the constructed visual measures are associated with audience engagement. If the measures capture attention-grabbing thumbnail characteristics, they should be positively related to video views, likes, and comments. Appendix Table A-2, Panel A, reports Pearson correlations between the visual measures and log(1+views), log(1+likes), and log(1+comments). Thumbnail Sentiment is positively correlated with all three engagement measures, and similar positive associations are observed for color saturation, contrast, face emotion intensity, and OCR sentiment. Although these associations are descriptive rather than causal, they suggest that the constructed visual measures capture salient thumbnail characteristics associated with higher audience engagement.

Second, we compare videos in the bottom and top deciles of the Thumbnail Sentiment distribution. Appendix Table A-2, Panel B, summarizes low- and high-sentiment thumbnail deciles together with their OCR text and component scores. Relative to the bottom decile, top-decile thumbnails exhibit substantially higher color saturation, contrast, face-expression intensity, and OCR sentiment. These differences suggest that Thumbnail Sentiment captures attention-grabbing visual characteristics. Together, the tail comparisons provide construct validity for the Thumbnail Sentiment measure.

3.4 Transcript Sentiment, Linguistic Complexity, and Firm-Level Background Tone

The spoken narrative conveys substantive interpretations of firm fundamentals, market developments, and strategic commentary. To characterize emotional and informational properties of transcripts, we apply two complementary sentiment frameworks.

First, affective polarity is measured using VADER, a sentiment analyzer calibrated for online conversational English and informal digital communication (Hutto and Gilbert, 2014).

Segment-level VADER scores are aggregated to produce video-level sentiment. To capture systematic differences in how creators discuss specific firms, we compute a Firm-Level Background Affective Tone as the historical average of VADER sentiment for each firm across all its video appearances.

Second, financial sentiment is measured using the Loughran–McDonald (LM) dictionary, which provides a domain-specific lexicon for positive and negative financial terms (Loughran and McDonald, 2011). Both video-level LM scores and Firm-Level Background LM Tone are computed to isolate incremental information conveyed by each individual video.

Linguistic richness is quantified using natural-language-processing metrics including lexical diversity, syntactic depth, readability indices, vocabulary density, and entropy. Textual complexity has been shown to influence information processing costs and investor responses (Li, 2008; Bloomfield, 2002), as well as the perceived credibility of online financial information (Paydarzarnaghi et al., 2024). These metrics capture cognitive load and potential belief dispersion arising from complex communication styles.

3.5 Linking YouTube Events to CRSP and Compustat

Appendix Table A-3 describes the construction of five video content measures constructed from principal components analysis (PCA), and Appendix Table A-4 reports the corresponding first-principal-component loadings. Each video content score is constructed at the video level from standardized inputs and then assigned to each associated video–firm event. Appendix Table A-5 reports descriptive statistics for selected underlying components. The regression specifications use standardized versions of the composite scores.

To examine price and volume reactions, each YouTube–Firm Event is linked to a publicly traded security using the CRSP–Compustat Merged (CCM) Permanent Link Table, which provides stable mappings across ticker changes, restructurings, and share-class adjustments. Each firm mention is mapped to a Compustat Ticker (TIC), then to a Compustat Firm Identifier (GVKEY), and finally to a CRSP Security Identifier (PERMNO). This mapping ensures consistency in the presence of ticker changes and firm reorganizations.

Event timing is aligned to the CRSP trading calendar. Videos published before the market opens are assigned to that trading day, while videos published after market close, on weekends, or on holidays are assigned to the next available trading day. Videos published during regular trading hours are assigned to the same calendar trading day for purposes of defining $t = 0$. Because daily CRSP outcomes aggregate the entire trading day, the $t = 0$ outcome for intraday videos may include trading that occurred before the video was released. Accordingly, we report $t = 0$ coefficients only as contemporaneous benchmark associations. The main analysis of market reaction relies on horizons that begin at $t + 1$, including $CAR[t + 1, t + 5]$, $CAR[t + 1, t + 22]$, $CAR[t + 1, t + 66]$, and the corresponding cumulative volume and volatility measures. Throughout the interpretation, we distinguish $t = 0$ benchmark estimates (used to gauge the extent of subsequent reversals) from post-release response windows. The latter begin at $t + 1$ and provide the basis for our main inferences about market reactions after video release.

The resulting dataset is a consolidated event file linking the Video Event Identifier, Firm Ticker, GVKEY, PERMNO, and aligned Market Event Date.

3.6 Integration with Compustat Fundamentals

Firm characteristics are merged from Compustat Quarterly data. To avoid look-ahead bias, an as-of merge is performed: for each event date, the most recent quarterly filing with a fiscal date not exceeding the event date is selected, consistent with event-study methodology emphasizing the importance of restricting the information set to that available at the time of the event (MacKinlay, 1997). The resulting dataset includes firm size, book-to-market ratios, leverage, profitability, operating income, firm age, and industry classification.

3.7 CRSP Market Data and Construction of Abnormal Returns, Trading Activity, and Volatility

Daily CRSP data provide price, return, volume, and shares outstanding histories for each PERMNO. For every YouTube–Firm Event, we extract data for windows spanning $[-1, +66]$ days around the event date. Expected returns are computed using the Fama–French five-factor model, and Abnormal Returns are defined as the difference between realized and expected returns. Cumulative Abnormal Returns (CAR) summarize responses over different horizons (Brown and Warner, 1985).

Trading activity responses are captured through Abnormal Log Volume and cumulative abnormal volume (CAVOL), identifying deviations from typical liquidity patterns following information shocks.

Volatility responses are measured using abnormal squared returns (AR), with cumulative abnormal volatility (CAV) reflecting disagreement and uncertainty dynamics (Tetlock, 2007; Engelberg, 2008). Raw returns, cumulative raw returns, and raw volume/volatility series are computed for benchmarking.

Appendix Table A-7 provides formal definitions of the raw and abnormal return, volume, and volatility measures, and Appendix Table A-8 defines the event windows used throughout the analysis.

3.8 Construction of the Final Event-Study Panel

The final dataset integrates all multimodal components into a unified event-level panel, with each row representing a distinct YouTube–Firm Event. Each observation includes (i) granular YouTube-derived features (visual salience, textual sentiment, linguistic complexity, creator reputation); (ii) Compustat firm attributes; and (iii) CRSP-based return, volume, and volatility reactions across multiple windows.

This pipeline—from raw video collection, Whisper transcription, sentiment and complexity extraction, entity resolution, event-date alignment, and CRSP–Compustat merging—produces a dataset suited for analyzing how video financial communication affects asset prices, trading behavior, and belief dispersion. Appendix Table A-9 reports descriptive statistics for the event-level dependent variables (Panel A), and PCA-based content measures (Panel B) used as independent variables.

3.9 Baseline regression model

To examine how YouTube video characteristics influence subsequent market reactions, we estimate the following baseline panel regression model:

$$\begin{aligned} Y_{i,t+h} = \alpha &+ \beta_1 \, \text{Thumbnail Sentiment}_{i,t} + \beta_2 \, \text{Title Sentiment}_{i,t} \\ &+ \beta_3 \, \text{Transcript Sentiment}_{i,t} + \beta_4 \, \text{Text Richness}_{i,t} \\ &+ \beta_5 \, \text{Influencer Reputation}_{i,t} + \gamma' X_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t+h} \end{aligned} \tag{1}$$

where i denotes firms, t is the YouTube video release date, and h represents the forward-looking market horizon. The dependent variable $Y_{i,t+h}$ captures multiple dimensions of market response, including daily abnormal returns ($AR_{i,t+h}$), cumulative abnormal returns (CAR), abnormal trading volume (AVOL and CAVOL), and abnormal volatility (AV and CAV), as defined in Subsection 3.7.

To capture both immediate and delayed effects, we construct cumulative abnormal returns over the horizons $[t+1, t+5]$, $[t+1, t+22]$, and $[t+1, t+66]$, corresponding respectively to the first trading week, the first trading month, and the first trading quarter following the video's release. All event windows are defined in trading days, consistent with standard event-study methodology. These multiple horizons allow us to assess whether content-driven effects dissipate quickly, persist over time, or exhibit delayed adjustment consistent with attention cycles or gradual information diffusion. The main explanatory variables—Thumbnail Sentiment, Title Sentiment, Transcript Sentiment, Text Richness, and Influencer Reputation—are standardized to have mean zero and unit variance to facilitate economic interpretation and comparability across coefficients.

The vector of control variables, $X_{i,t}$, includes firm characteristics commonly used in the asset-pricing and corporate-disclosure literatures, including firm size (log market capitalization), book-to-market ratio, leverage, past return volatility, and recent stock performance (momentum). We additionally incorporate firm-level Fama–French five-factor (FF5) betas estimated from a 60-month rolling window to ensure that our results are not driven by exposure to standard risk factors.

To absorb unobservable heterogeneity, all regressions include firm fixed effects (μ_i), which account for time-invariant differences across firms (such as disclosure style, investor base, or information environment), and time fixed effects (λ_t), implemented as year-by-month fixed effects to capture common shocks, macroeconomic conditions, and systematic shifts in investor attention. Standard errors are clustered at the firm level to address serial correlation in firm-specific observations.

This specification isolates within-firm variation in YouTube video characteristics and provides a framework for evaluating how multimodal content influences short- and medium-horizon price dynamics, trading activity, and volatility.

Control Variables

The vector $X_{i,t}$ includes firm-level characteristics widely used in asset pricing and corporate finance. Specifically, we control for book-to-market ratio, cash holdings, firm age, leverage, log market capitalization (firm size), log stock price, operating profitability, return on assets, sales growth, 60-day turnover, and 60-day realized volatility. These controls absorb known determinants of returns, volatility, and trading activity (e.g., Fama and French, 1992; Novy-Marx, 2013; Hou et al., 2015), ensuring that the effects of video content are not driven by firm fundamentals or standard risk factors. Appendix Table A-9, Panel C, reports descriptive statistics for the firm-level controls.

Fixed Effects and Error Structure

We include firm fixed effects (μ_i) to absorb time-invariant heterogeneity across firms and month fixed effects (λ_t) to capture market-wide and macroeconomic conditions. Standard errors are clustered at the stock level to allow for arbitrary heteroskedasticity and serial correlation, following best practices for panel event-study designs (Hirshleifer et al., 2009).

4 Empirical Results

4.1 Baseline Results: Raw Returns

We begin with a baseline specification using raw returns, which provides a reduced form view of how markets respond to multimodal YouTube content prior to adjusting for systematic risk exposures or other potential confounding factors. This approach is standard in empirical asset pricing (Fama and French, 1993; Daniel and Titman, 1997) and is useful for documenting unconditional price patterns, though it does not by itself establish causal effects.

More broadly, this baseline analysis provides an initial characterization of how user-generated, multimodal digital content maps into asset price movements before disentangling attention-driven reactions from risk-adjusted informational effects. Although the regressions include standard controls and fixed effects, the raw-return specification in Table 2 should still be interpreted primarily as descriptive rather than causal evidence.

4.2 Visual Cues and Immediate Price Pressure

The raw-return results in Table 2 indicate that visual sentiment embedded in thumbnails generates economically meaningful contemporaneous price responses. Thumbnail sentiment is positively associated with returns at $t = 0$ (0.016, $p < 0.05$), but this effect reverses sharply at $t + 1$ (-0.015, $p < 0.05$) and becomes increasingly negative over longer horizons (e.g.,

-0.069 at $t + 22$, $p < 0.01$).

This pattern is consistent with attention-based trading models in which salient visual stimuli attract investor attention and induce temporary buying pressure, followed by subsequent correction (Da et al., 2011). The rapid reversal suggests that these visual signals are not strongly anchored in fundamentals but instead reflect short-lived attention shocks.

Title sentiment exhibits a related but less consistent pattern. While title sentiment is positively associated with same-day returns (0.024, $p < 0.01$), its effects are less stable across horizons, with mixed signs and limited statistical significance beyond the immediate window. This suggests that, although titles contribute to attention capture, their impact on pricing dynamics is weaker and more transient than that of thumbnails.

4.3 Textual Sentiment and Short-Run Reversals

In the raw-return specification, textual sentiment displays a reversal pattern rather than the slower diffusion process often associated with informational tone. Positive transcript tone is associated with higher returns at $t = 0$ (0.030, $p < 0.01$), but this effect reverses in the short run, with negative coefficients at $t + 1$ (-0.016, $p < 0.01$) and $t + 5$ (-0.032, $p < 0.01$), and becomes statistically insignificant at longer horizons.

This pattern suggests that, in raw returns, textual sentiment may also capture short-term attention-driven trading or correlated firm characteristics rather than purely fundamental information. Because raw returns embed systematic risk exposures and other confounding factors, they do not cleanly isolate informational effects. We revisit this distinction in the abnormal-return analysis, where risk adjustment allows for a clearer separation between attention-driven price pressure and information-based valuation effects (e.g., Hirshleifer and Teoh, 2003; Tetlock et al., 2008).

4.4 Linguistic Richness and Valuation Discounts

Textual richness is generally associated with negative returns, with coefficients ranging from approximately -0.029 to -0.108. The estimates are statistically significant in most short- and long-horizon specifications, although not uniformly across all intervals.

This consistent negative relationship is in line with theories linking linguistic complexity to higher processing costs and greater perceived uncertainty (Miller, 2010). One interpretation is that more complex or dense content increases ambiguity, leading investors to apply higher discount rates. However, because raw returns may also reflect correlated firm characteristics, this relationship should be interpreted cautiously as reduced-form evidence.

4.5 Influencer Reputation and Overreaction Dynamics

Influencer reputation is positively associated with contemporaneous returns (0.023, $p < 0.01$) but becomes negative at longer horizons (e.g., -0.024 at $t+22$ and -0.028 at $t+66$), although these later effects are not always statistically significant.

This pattern is consistent with an overreaction–correction dynamic in which high-visibility communicators attract initial investor attention that partially reverses over time (Barber et al., 2009). At the same time, this relationship may reflect selection effects, as prominent influencers may disproportionately cover firms that are already salient or subject to higher uncertainty.

4.6 Economic Interpretation

Taken together, the raw-return results reveal three broad empirical patterns. First, visual attributes—particularly thumbnails—are associated with short-term price pressure that reverses quickly. Second, textual sentiment exhibits short-run reversal dynamics rather than clear long-run information diffusion in raw returns.

Third, linguistic richness is generally negatively related to returns, suggesting a role for processing frictions or perceived uncertainty. These patterns indicate that multimodal video content affects markets through a combination of attention and complexity channels. However, because raw returns incorporate both firm-specific information and systematic risk exposures, these findings should be interpreted as descriptive rather than causal.

4.7 Motivation for Abnormal-Return Analysis

Because raw returns embed exposures to size, value, momentum, and volatility factors, they do not isolate content-driven pricing effects. The next step therefore examines abnormal returns (AR/CAR), which adjust for systematic risk and provide a cleaner measure of content-related market reactions.

This transition is essential for distinguishing whether the observed patterns reflect genuine information effects or simply correlated risk exposures and omitted variables. By moving to a risk-adjusted framework, we can more precisely assess the extent to which multimodal video content contributes to price discovery versus temporary market distortions.

5 Main Regression Results

The central empirical question is how multimodal YouTube disclosures shape market reactions through attention, information, complexity, and reputation channels. Across returns, volatility, and trading activity, the estimates reveal a heterogeneous structure. First, visual attributes primarily generate attention-driven responses, particularly in trading activity and short-run price movements. Second, transcript sentiment is more closely related to return dynamics, although its effects are weaker and less stable across specifications. Third, textual richness exhibits strong short-run effects in trading activity but only limited and economically small effects in returns, with no clear long-run pattern. Fourth, influencer reputation amplifies investor participation and is associated with increased volatility over time, consistent with disagreement-driven uncertainty, as higher volatility is commonly interpreted as reflecting increased dispersion in investor beliefs.

5.1 Abnormal Returns and Market Reactions

A necessary first step in interpreting return reactions is to verify that pre-event dynamics are muted. Placebo tests using pre-event windows (reported in Appendix Tables) indicate that coefficients on thumbnail, title, transcript, and textual richness variables are economically small and often statistically insignificant. This muted pre-trading response is consistent with limited anticipatory positioning and supports the credibility of the event-time identification strategy. Following standard practice in event-time designs (Hirshleifer et al., 2009), weak pre-event coefficients reduce concerns about reverse causality and anticipatory behavior.

Turning to post-event dynamics, the results in Table 3 indicate heterogeneous responses across content channels.

Thumbnail sentiment produces a negative contemporaneous return at $t = 0$ (-0.012, $p < 0.01$), with effects remaining negative but attenuated over subsequent horizons, becoming economically small by $t + 66$ (-0.005, $p < 0.1$). This pattern is consistent with short-lived price pressure driven by visual salience, in line with attention-based mispricing frameworks (Barber and Odean, 2008).

Title sentiment exhibits a reversal dynamic. While the contemporaneous effect is negligible, returns decline in the short run (-0.010 at $t + 1$, $p < 0.01$) before turning positive at longer horizons (0.013 at $t + 66$, $p < 0.01$). This intertemporal pattern is consistent with delayed price adjustment, although the magnitude and timing of the reversal vary across horizons. and it is consistent with headline-salience effects documented in (Hirshleifer et al., 2009).

Transcript sentiment behaves differently from visual features. It generates a positive contemporaneous return (0.012, $p < 0.01$), but the effect dissipates over time, becoming small and only weakly significant by $t + 66$ (0.005, $p < 0.1$). This attenuation suggests that transcript sentiment influences initial price formation but does not generate strong or persistent return predictability over longer horizons.

Textual richness exhibits limited short-run effects and no clear long-run pattern, with the coefficient remaining economically small at $t + 66$ (-0.003). This muted response provides only limited support for processing-friction mechanisms (Li, 2008; You and Zhang, 2009).

Finally, influencer reputation produces a positive contemporaneous effect (0.010, $p < 0.01$) but is followed by a significant long-run reversal (-0.016 at $t + 66$, $p < 0.01$). This asymmetric pattern suggests that reputation initially attracts investor attention but does not translate into persistent valuation effects, consistent with intermediary-based attention channels.

Overall, the return evidence indicates that video content characteristics generate short-lived market reactions but do not lead to sustained abnormal returns. This pattern is consistent with attention-driven trading models in which salient information triggers temporary price pressure without persistent mispricing (Barber and Odean, 2008; Bordalo et al., 2020).

5.2 Abnormal Volatility

As shown in Table 4, the pre-event window provides a diagnostic check for anticipatory uncertainty. Coefficients in the pre-event (placebo) windows are economically small and statistically weak across all multimodal attributes, indicating no systematic increase in volatility prior to the video release. This absence of anticipatory uncertainty supports the validity of the event-time design and suggests that the observed volatility responses reflect post-disclosure reactions rather than pre-positioning (Kandel and Pearson, 1995; Harris and Raviv, 1993) .

Post-event dynamics, however, reveal distinctions across channels. Thumbnail sentiment does not exhibit a meaningful contemporaneous effect but becomes positive at longer horizons (0.008 at $t + 66$, $p < 0.05$). Title sentiment exhibits a negative short-run effect (-0.156 at $t + 1$, $p < 0.01$) and remains economically small thereafter, with no strong long-run persistence (0.0006 at $t + 66$). Because visual cues heighten salience without necessarily adding information, these mixed volatility responses provide only limited support for H1.

Transcript sentiment exhibits weak and statistically insignificant effects across horizons, with no clear evidence of sustained volatility reduction (e.g., 0.002 at $t + 66$). This provides limited support for information-clarification effects (Tetlock, 2007; Cookson and Niessner, 2020) and suggests that linguistic tone does not materially alter uncertainty dynamics.

Textual richness produces a consistent decline in volatility, with a large contemporaneous effect (-0.146, $p < 0.01$) and persistent negative effects through longer horizons (-0.031 at $t + 66$, $p < 0.01$). This pattern is more consistent with processing-friction mechanisms that dampen immediate market reactions (Callen et al., 2013) rather than amplify disagreement, providing mixed evidence for H3.

Influencer reputation, in contrast, increases volatility over time, becoming significant in the short run (0.002, $p < 0.01$) and rising to 0.027 ($p < 0.01$) by $t+66$. This pattern suggests increasing disagreement rather than stabilization, providing partial support for H4.

5.3 Abnormal Trading Volume

The $t - 1$ pre-event window again shows minimal activity across all multimodal attributes. Coefficients are economically negligible, indicating no anticipatory trading. This clean pre-event profile reinforces that the volume effects reflect genuine post-disclosure reactions rather

Post-event responses are substantial. In Table 5, thumbnail sentiment exhibits a negative contemporaneous effect at $t = 0$ (-0.035, $p < 0.01$), followed by a reversal at $t+1$ (0.023, $p < 0.10$) and a modest cumulative increase by $t+66$ (0.006, $p < 0.10$). Title sentiment generates strong and persistent trading responses, with a significant effect at $t = 0$ (0.038, $p < 0.01$) and continued increases through $t + 66$ (0.016, $p < 0.01$). These patterns broadly align with attention-cycle dynamics documented in (Da et al., 2011; Cookson and Niessner, 2020), providing support for H1, although the initial negative response for thumbnails suggests short-run heterogeneity in attention allocation.

Transcript sentiment produces limited and short-lived volume effects, with a negative coefficient at $t+1$ (-0.022, $p < 0.05$) and no persistent impact at longer horizons. This muted response is consistent with evidence that information-driven tone primarily affects prices rather than trading intensity (Chen et al., 2014; Hassan et al., 2019), offering only weak support for H2.

Textual richness generates strong contemporaneous and short-run trading responses, with coefficients of 0.065 ($p < 0.01$) at $t = 0$ and 0.052 ($p < 0.01$) at $t + 1$, but these effects dissipate at longer horizons and become statistically insignificant by $t + 66$. This pattern is consistent with speculative trading induced by interpretive complexity (Diether et al., 2002; Scheinkman and Xiong, 2003), providing partial—rather than strong— support for H3.

Influencer reputation produces a positive and persistent effect on trading activity, with significant increases at $t = 0$ (0.056, $p < 0.01$), $t + 1$ (0.042, $p < 0.01$), and continuing through longer horizons (0.016, $p < 0.01$ at $t + 66$). This pattern suggests sustained investor engagement consistent with attention amplification and investor participation channels (Fang and Peress, 2009), strengthening support for H4 through an attention-based rather than a purely transient mechanism.

5.4 Integrated Interpretation and Hypothesis Assessment

Taken together, the evidence highlights a heterogeneous pattern of market responses to multimodal YouTube content. Visual cues primarily operate through attention channels, generating strong short-run trading responses, while thumbnail effects exhibit reversals and limited long-run persistence, providing partial support for H1.

Transcript sentiment plays a limited role in trading activity and exhibits modest and short-lived effects, reinforcing its interpretation as an informational channel primarily affecting returns rather than volume, and thus offering weak support for H2.

Textual richness generates pronounced short-run trading responses that dissipate over time, while its volatility effects suggest dampening rather than amplification of disagreement, yielding mixed support for H3.

Influencer reputation produces sustained increases in trading activity, consistent with attention amplification and broader investor participation. However, its effects vary across outcomes and horizons, providing partial rather than uniform support for H4.

Overall, these findings indicate that multimodal digital disclosures influence markets through multiple interacting channels, with effects that vary across outcomes and time horizons rather than following a single uniform mechanism.

5.5 Control Variables and Model Specification

Control variables behave in theoretically consistent ways across all specifications. Book-to-market predicts negative returns and higher volatility (Fama and French, 1993). Leverage increases risk exposure (French et al., 1987). Firm size is associated with lower information asymmetry and deeper liquidity, which contribute to lower volatility and more efficient trading dynamics (Amihud and Mendelson, 1986; Lang and Lundholm, 1996). Momentum retains strong predictive power for returns and trading activity (Jegadeesh and Titman, 1993). Turnover is highly persistent, reflecting attention and liquidity cycles (Chordia et al., 2001).

All regressions include firm fixed effects to absorb time-invariant heterogeneity in disclosure style, investor base, and information environment. Standard errors are clustered at the stock level. Robustness tests using alternative control lags, industry-by-month fixed effects, and alternative volatility and volume windows yield stable estimates, indicating that multimodal effects are not artifacts of omitted variables or correlated risk exposures.

6 Robustness Tests

This section evaluates whether the main results are sensitive to alternative empirical choices or to pre-existing trends around video releases. We consider three sets of tests. First, we reestimate the main specifications using alternative event windows. Second, we use propensity score matching to compare high- and low-content-intensity observations with similar observable characteristics. Third, we examine pre-event placebo windows and randomized content assignments. These tests are intended to assess the stability of the associations documented above; they do not eliminate all concerns related to omitted information events or endogenous video timing.

6.1 Alternative Event Windows

We first examine whether the results depend on the choice of event window. In the main analysis, we study contemporaneous responses and cumulative outcomes over short and medium horizons. As an alternative, Tables A-10–A-12 reestimate the return, volatility, and trading-volume regressions using the post-event windows $[t+4, t+15]$ and $[t+4, t+30]$. These windows begin after the immediate post-release period and therefore help distinguish short-run attention effects from more persistent market responses. The use of alternative event windows follows standard event-study practice (Fama et al., 1969; MacKinlay, 1997; Kothari and Warner, 2007).

The return results in Table A-10 are broadly consistent with the baseline evidence. Thumbnail sentiment produces a negative contemporaneous response similar in magnitude to the main specification, but the effect becomes smaller over medium horizons. Title sentiment also displays reversal patterns, although the timing and magnitude vary across windows. Transcript sentiment shows some positive short-run response, but the estimates are less stable across horizons. Textual richness and influencer reputation also produce non-monotonic patterns, suggesting that these variables are associated with market reactions that depend on the timing of measurement. Overall, the alternative-window return results support the main conclusion that video characteristics are associated with short-lived price movements rather than persistent abnormal returns.

The volatility results in Table A-11 also indicate heterogeneous responses across content channels. Thumbnail and title sentiment are associated with higher volatility in several post-event windows, consistent with the view that visually salient or emotionally framed content may coincide with greater investor disagreement (Harris and Raviv, 1993; Kandel and Pearson, 1995). Transcript sentiment has weaker and less consistent effects. Textual richness is generally associated with lower volatility in these specifications, which provides mixed evidence for the linguistic-complexity hypothesis. This pattern suggests that richer language may sometimes dampen immediate disagreement, perhaps because more detailed discussion attracts a different set of viewers or provides additional context. Influencer reputation is positively related to volatility in several medium-horizon windows, consistent with broader audience reach and more dispersed interpretation.

The trading-volume results in Table A-12 provide the clearest support for an attention-based interpretation. Thumbnail and title sentiment are associated with increases in trading activity in the early post-event period, although the effects are not uniform across all windows. This pattern is consistent with attention-driven trading models in which salient information attracts investor participation but need not generate lasting price effects (Barber and Odean, 2008; Da et al., 2011; Rakowski et al., 2021). Transcript sentiment produces more modest volume responses, while textual richness and influencer reputation display horizon-specific patterns. Taken together, the alternative-window tests indicate that the main trading-volume results are not confined to a single event-window definition, but they also show that the timing of the response matters.

Overall, the alternative-window evidence supports the main interpretation that multimodal YouTube features are associated with short-run market activity, especially trading volume and volatility. The results are less consistent for longer-horizon returns, which cautions against interpreting the evidence as persistent mispricing or durable changes in firm value.

6.2 Matched-Sample Tests

We next use propensity score matching to assess whether the baseline results are driven by observable differences across firms or videos. For each main content dimension—Thumbnail Sentiment, Title Sentiment, Transcript Sentiment, Text Richness, and Influencer Reputation—we classify observations in the top quartile as treated and observations in the bottom quartile as controls. Intermediate observations are excluded to sharpen the treatment contrast. Propensity scores are estimated using firm characteristics and non-focal content variables, and observations are matched using one-to-one nearest-neighbor matching with a caliper restriction (Rosenbaum and Rubin, 1983; Austin, 2011; Imbens and Rubin, 2015). We then reestimate the baseline regressions on the matched sample with firm and year–month fixed effects. Covariate balance is assessed using standardized mean differences (Stuart, 2010).

The matched-sample return results in Table A-13 show economically small effects across most specifications. A few coefficients are statistically different from zero, but the overall pattern does not suggest systematic abnormal return predictability. This evidence is consistent with the baseline conclusion that video features are more closely associated with attention and trading activity than with persistent valuation effects.

The matched-sample volatility results in Table A-14 provide somewhat stronger evidence for disagreement-related responses. Influencer reputation is positively associated with volatility over longer horizons, while other content measures display weaker or less consistent effects. This finding is consistent with the idea that widely followed creators may broaden the audience for a stock-specific video and thereby increase heterogeneity in interpretation. However, because matching only balances observable characteristics, these results should be interpreted as matched-sample associations rather than causal estimates.

The matched-sample trading-volume results in Table A-15 are also consistent with the main evidence. Title tone and influencer reputation are associated with higher trading activity around the event date, while transcript-based measures are less consistently related to volume. These results reinforce the interpretation that easily observed presentation features and creator reach are associated with investor participation. They also suggest that the main volume results are not solely attributable to observable differences between firms receiving high- and low-intensity YouTube coverage.

Taken together, the matched-sample tests support the central interpretation of the paper. Video characteristics are not strongly related to persistent abnormal returns, but several presentation and creator-level measures are associated with trading activity and volatility. This pattern is consistent with attention-based trading and heterogeneous investor interpretation (Barber and Odean, 2008; Cookson and Niessner, 2020), while remaining cautious about causal interpretation.

6.3 Pre-Event Placebo Tests

Finally, we examine whether the main results are preceded by similar patterns before video release. We estimate the baseline models over five non-overlapping pre-event windows: $[t-66, t-53]$, $[t-52, t-39]$, $[t-38, t-25]$, $[t-24, t-11]$, and $[t-10, t-1]$. If the documented post-event associations were driven primarily by persistent pre-trends, one would expect similar patterns in these placebo windows. We also conduct a randomized placebo test in which key video-content variables are reassigned across observations, breaking the link between the video characteristics and the corresponding firm-event outcomes.

Tables A16–A18 report the placebo results for abnormal returns, abnormal volatility, and abnormal trading volume. Across outcomes, the coefficients on the main video-content variables are generally small and statistically insignificant. A limited number of coefficients reach conventional significance levels, but these effects are not stable across adjacent pre-event windows and do not form a clear economic pattern. The randomized placebo tests produce similarly weak and unsystematic results.

These placebo tests reduce concerns that the main findings are mechanically driven by pre-existing trends in returns, volatility, or trading volume. They also suggest that the market reactions documented in the main analysis are concentrated after video release rather than appearing well before the event. At the same time, the placebo tests cannot rule out all omitted information events or unobserved shocks that occur close to the release date. Accordingly, the evidence should be interpreted as supporting the event-time association between video characteristics and market reactions, rather than proving that video content alone causes the observed outcomes.

6.4 Summary of Robustness Evidence

The robustness tests lead to three conclusions. First, the results are not driven by a single event-window definition. The timing and magnitude of the estimates vary, but the main evidence of short-run trading and volatility responses remains visible across alternative windows. Second, the matched-sample tests indicate that the results are not explained solely by observable differences between firms or videos with high versus low content intensity. Third, the placebo tests show little evidence of systematic pre-event patterns. In untabulated analysis, we also confirm that our results are robust to alternative clustering methods.

Overall, the additional tests support a conservative interpretation of the paper's main findings. Multimodal YouTube features—especially visual and title-based cues and creator reputation—are associated with trading activity and volatility around video releases. The return evidence is weaker and less persistent. This pattern is consistent with investor-attention and belief-dispersion channels, but it does not establish that YouTube videos generate lasting mispricing or that the estimated relations are fully causal.

7 Conclusion

This paper examines whether the visual and linguistic features of finance-related YouTube videos are associated with firm-level market reactions. Using an event-time panel of videos from 25 finance-focused YouTube channels from 2021 to 2024, we link video releases to CRSP and Compustat data and study abnormal returns, abnormal trading volume, and abnormal volatility around the first trading day on which investors could respond to the content. The empirical design does not require video releases to be exogenous. Instead, the analysis asks whether different features of video content are associated with differential post-release market reactions after controlling for firm characteristics, firm fixed effects, time fixed effects, and other video attributes.

The evidence suggests that visual and title-based features are most closely related to investor attention. Videos with more salient thumbnails or emotionally framed titles are generally followed by higher trading activity and, in some specifications, higher volatility. Return effects are more limited and tend to be short-lived, often weakening or reversing across longer horizons. This pattern is consistent with temporary attention-based trading pressure rather than a persistent change in firm value. These findings extend evidence that investor attention affects trading and short-run price pressure (Barber and Odean, 2008; Da et al., 2011; Rakowski et al., 2021) by showing that attention may be shaped by the visual presentation of financial information, not only by news coverage, search activity, or textual tone.

The results also refine how attention-based theories should be applied to digital financial media. In traditional attention models, investors respond to salient stocks, news events, or information sources because attention is scarce and search costs are high (Barber et al., 2009; Hirshleifer and Teoh, 2003; Hirshleifer et al., 2009). Financial YouTube introduces an additional layer: the form of presentation itself may become part of the attention shock. Thumbnails, titles, facial expressions, and visual intensity are not merely delivery mechanisms for information already contained in the transcript. They are observable signals that can affect whether investors attend to the content and how they frame the underlying message. This suggests that models of attention in retail-dominated information environments may need to distinguish between informational salience and presentational salience.

Transcript sentiment plays a more limited role. Although positive spoken tone is associated with some short-run return responses, the effects are weaker and less stable across specifications than those for visual and title-based measures. This evidence suggests that spoken content and visual presentation are not substitutes. Rather, financial video appears to combine informational content with presentation features that affect how investors allocate attention. This distinction connects the paper to the media-sentiment literature, which shows that financial language can predict market reactions (Tetlock, 2007; Tetlock et al., 2008; Loughran and McDonald, 2016), while suggesting that video-based communication requires a broader measurement framework than text alone.

Textual richness also produces mixed evidence. It is associated with short-run trading responses, but its relation with returns and volatility varies across horizons. These findings suggest that linguistic complexity may affect investor processing, but the direction and persistence of that effect depend on the outcome and event window considered. In this respect, the evidence is consistent with the broader view that complex information can slow processing or increase heterogeneity in interpretation (Bloomfield, 2002; Li, 2008; Miller, 2010), but it does not imply that richer language uniformly increases disagreement or volatility. For theory, this suggests that processing frictions may operate differently in informal digital media than in formal disclosures. In YouTube videos, complexity may sometimes reflect deeper discussion, but it may also filter out less attentive viewers or shift reactions toward investors with greater information-processing ability.

Creator-level visibility also matters. Measures of influencer reputation are associated with higher trading participation and, in some specifications, higher volatility. These findings are consistent with the view that widely followed creators can broaden the audience for firm-specific information. At the same time, the evidence does not imply that reputation uniformly improves price discovery or that high-reputation creators systematically move prices. Rather, creator reach appears to condition how strongly investors respond to the same broad type of financial content. This result connects YouTube creators to the broader literature on media intermediaries, social transmission, and disagreement (Fang and Peress, 2009; Cookson and Niessner, 2020; Hirshleifer et al., 2025). It also suggests that informal digital intermediaries may affect markets not only through what they say, but through the size and composition of the audience they mobilize.

Several features of the evidence warrant caution. The analysis is observational, and financial YouTube videos may be released in response to firm news, market movements, or investor interest that is not fully captured by the controls. Placebo tests using pre-event windows reduce concerns about simple pre-trends, and the results are generally similar across alternative windows and matched-sample specifications. However, these tests do not eliminate the possibility that omitted information events or unobserved attention shocks partly explain the results. The findings should therefore be interpreted as evidence that video characteristics are associated with market reactions, not as definitive causal estimates of the effect of YouTube content.

The study also has measurement limitations. Transcripts generated from audio may contain errors in firm names, tickers, or technical terminology. OCR-based extraction of thumbnail text may be noisy when fonts are stylized or images are crowded. Firm matching based on tickers, names, and semantic similarity may miss ambiguous references or incorrectly assign firms in some cases. In addition, the sample is limited to U.S. equities, English-language content, and a selected set of large finance-focused YouTube channels. The results may not generalize to smaller creators, non-U.S. markets, other languages, or other video platforms.

Despite these limitations, the findings show that multimodal digital media are relevant for understanding investor attention and market activity. Financial YouTube content is not simply text delivered through another medium. It combines spoken analysis, visual design, title framing, and creator reputation. The evidence suggests that these features are associated with trading activity, volatility, and short-run price movements in ways that are consistent with attention-based trading and heterogeneous investor interpretation. Future research could examine whether these effects are stronger among stocks with high retail ownership, low institutional ownership, low analyst coverage, low liquidity, or high social-media attention. Additional work could also validate visual sentiment measures through hand coding or investor experiments, and could compare YouTube with other platforms where financial information is increasingly shaped by algorithmic ranking and creator-driven presentation.

References

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

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223–249.

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, 10(2):150–161.

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

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

Bloomfield, R. J. (2002). The incomplete revelation hypothesis and financial reporting. *Accounting Horizons*, 16(3):233–243.

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

Brown, S. J. and Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of financial economics*, 14(1):3–31.

Callen, J. L., Khan, M., and Lu, H. (2013). Accounting quality, stock price delay, and future stock returns. *Contemporary Accounting Research*, 30(1):269–295.

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

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *Journal of Finance*, 56(2):501–530.

Cookson, J. A. and Niessner, M. (2020). Why don't we agree? evidence from a social network of investors. *Journal of Finance*, 75(1):173–228.

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

Daniel, K. and Titman, S. (1997). Evidence on the characteristics of cross-sectional variation in stock returns. *Journal of Finance*, 52(1):1–33.

Diether, K. B., Malloy, C. J., and Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *Journal of Finance*, 57(5):2113–2141.

Engelberg, J. (2008). Costly information processing: Evidence from earnings announcements. AFA 2009 San Francisco Meetings Paper.

Fama, E. F., Fisher, L., Jensen, M. C., and Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, 10(1):1–21.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.

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

Fang, L. and Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, 64(5):2023–2052.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.

Harris, M. and Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial Studies*, 6(3):473–506.

Hassan, T. A., Hollander, S., van Lent, L., and Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, 134(4):2135–2202.

Hirshleifer, D., Lim, S. S., and Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance*, 64(5):2289–2325.

Hirshleifer, D., Peng, L., and Wang, Q. (2025). News diffusion in social networks and stock market reactions. *Review of Financial Studies*, 38(3):883–937.

Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1–3):337–386.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705.

Hutto, C. J. and Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 8, pages 216–225.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press, Cambridge.

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

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.

Kandel, E. and Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872.

Koh, B. and Cui, F. (2022). An exploration of the relation between the visual attributes of thumbnails and the view-through of videos: The case of branded video content. *Decision Support Systems*, 160:113820.

Kothari, S. P. and Warner, J. B. (2007). Econometrics of event studies. In Eckbo, B. E., editor, *Handbook of Corporate Finance: Empirical Corporate Finance*, volume 1, chapter 1, pages 3–36. Elsevier.

Lang, M. H. and Lundholm, R. J. (1996). Corporate disclosure policy and analyst behavior. *Accounting Review*, 71(4):467–492.

Lang, P. J. (2000). Emotion and motivation: Attention, perception, and action. *Journal of Clinical Neurophysiology*, 17(5):451–456.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3):221–247.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1):13–39.

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85(6):2107–2143.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

Paydarzarnaghi, M. and Rakowski, D. (2026). Strategic bot activity in financial social media. *Finance Research Letters*, page 109808.

Paydarzarnaghi, M., Rakowski, D. A., and Yasar, M. (2024). Social media credibility and financial market activity. *Available at SSRN 5039937*.

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., and Sutskever, I. (2023). Robust speech recognition via large-scale weak supervision. In *Proceedings of the International Conference on Machine Learning*, volume 202, pages 28492–28518. PMLR.

Rakowski, D., Shirley, S. E., and Stark, J. R. (2021). Twitter activity, investor attention, and the diffusion of information. *Financial Management*, 50(1):3–46.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Scheinkman, J. A. and Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6):1183–1220.

Slovic, P., Finucane, M., Peters, E., and MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the affect heuristic for behavioral economics. *Journal of Socio-Economics*, 31(4):329–342.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1):1–21.

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

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, 63(3):1437–1467.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30.

You, H. and Zhang, X.-J. (2009). Financial reporting complexity and investor underreaction to 10-k information. *Review of Accounting Studies*, 14(4):559–586.

Table 1: Summary Statistics of YouTube video characteristics

Variable	N	Mean	Std. Dev.	P25	Median	P75
Video duration (minutes)	1,924	19.24	24.53	3.40	11.83	20.64
Views	1,924	325,925.91	1,776,689.19	830	7,166	71,205
Log(1 + views)	1,924	8.74	3.45	6.72	8.88	11.17
Likes	1,924	13,580.93	83,232.69	38	286	2,764
Log(1 + likes)	1,924	5.63	3.15	3.66	5.66	7.92
Comments	1,924	225.03	614.76	3	38	177
Log(1 + comments)	1,924	3.39	2.34	1.39	3.68	5.18
Number of firm mentions per video	1,924	229.63	395.08	22.00	112.00	273.25
Number of tickers matched per video	1,924	6.49	2.87	5.00	7.00	8.00
Indicator: posted during trading hours	1,924	0.279	0.449	0.000	0.000	1.000
Indicator: posted after market close	1,924	0.476	0.500	0.000	0.000	1.000
Indicator: weekend/holiday upload	1,924	0.245	0.430	0.000	0.000	0.000

Notes: This table reports summary statistics for the sample of 1,924 videos posted during 2021–2024. Views, likes, and comments are obtained from YouTube. The number of firm mentions and matched tickers is aggregated across all firms associated with a video. Upload-time indicators are mutually exclusive and are defined using U.S. Eastern Time. *Trading Hours* equals one for videos posted on non-holiday weekdays between 9:30 a.m. and 4:00 p.m.; *After Market Close* equals one for videos posted on non-holiday weekdays outside trading hours; and *Weekend/Holiday* equals one for videos posted on weekends or U.S. federal holidays. For indicator variables, the mean represents the sample proportion.

Table 2: Video Content Characteristics and Raw Returns (RR/CRR)

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment	0.016^{**}	-0.015^{**}	-0.019	-0.069^{***}	-0.056^{*}
	(0.006)	(0.007)	(0.014)	(0.024)	(0.032)
Title Sentiment	0.024^{***}	-0.003	0.029^{***}	-0.004	-0.014
	(0.005)	(0.005)	(0.011)	(0.017)	(0.023)
Text Sentiment	0.030^{***}	-0.016^{***}	-0.032^{***}	-0.001	-0.014
	(0.005)	(0.005)	(0.011)	(0.019)	(0.028)
Text Richness	-0.029^{***}	-0.018^{***}	-0.033^{***}	-0.008	-0.108^{***}
	(0.005)	(0.005)	(0.011)	(0.020)	(0.026)
Influencer Reputation	0.023^{***}	0.003	0.010	-0.024	-0.028
	(0.006)	(0.007)	(0.013)	(0.024)	(0.032)
Control Variables					
Book-to-Market	-0.301^{***}	-0.158^{***}	-0.913^{***}	-3.765^{***}	-11.917^{***}
	(0.037)	(0.038)	(0.085)	(0.175)	(0.282)
Leverage	0.198^{***}	-0.246^{***}	-0.510^{***}	-2.671^{***}	-7.787^{***}
	(0.022)	(0.023)	(0.045)	(0.082)	(0.124)
Log Market Cap	1.016^{***}	-0.832^{***}	-4.143^{***}	-18.800^{***}	-44.861^{***}
	(0.084)	(0.093)	(0.207)	(0.421)	(0.685)
Momentum 12M	-0.222^{***}	-0.095^{***}	-0.287^{***}	-0.039	0.042
	(0.011)	(0.012)	(0.024)	(0.043)	(0.060)
Turnover 60D	0.159^{***}	-0.016	-0.146^{***}	-0.387^{***}	1.579^{***}
	(0.015)	(0.016)	(0.032)	(0.065)	(0.094)
Volatility 60D	-0.092^{***}	0.046^{***}	-0.140^{***}	0.076	-1.078^{***}
	(0.013)	(0.013)	(0.027)	(0.053)	(0.074)
Within R^2 (%)	2.07	1.48	5.96	16.65	33.67
Adj. R^2 (%)	1.97	1.38	5.86	16.57	33.60
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	stock	stock	stock	stock	stock

Notes: This table presents estimates from regression analyses of video content characteristics on stock returns. Standard errors are clustered at the stock level. Analyses include firm, year-month fixed effects. All independent variables are standardized (log-transformed variables are not standardized). Thumbnail characteristics include all caps share, text color shares, and exclamation density. Content characteristics include numeric promises count, extreme terms count, total words, and all caps share. VIF check: All 5 PCA components have VIF ≤ 5 (acceptable). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Video Content Characteristics and Abnormal Returns (AR/CAR)

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	-0.012^{***}	-0.008^{*}	-0.004^{***}	-0.002	-0.005^{*}
	(0.004)	(0.004)	(0.001)	(0.002)	(0.003)
Title Sentiment Score	-0.001	-0.010^{***}	-0.0005	-0.007^{***}	0.013^{***}
	(0.003)	(0.003)	(0.0009)	(0.002)	(0.002)
Text Sentiment Score	0.012^{***}	0.003	-0.0005	-0.0010	0.005^{*}
	(0.003)	(0.003)	(0.0009)	(0.002)	(0.003)
Text Richness Score	-0.004	0.008^{**}	-0.002^{**}	0.006^{***}	-0.003
	(0.003)	(0.003)	(0.0010)	(0.002)	(0.003)
Influencer Reputation	0.010^{***}	-0.003	0.0007	-0.006^{***}	-0.016^{***}
	(0.004)	(0.004)	(0.001)	(0.002)	(0.003)
Control Variables					
Book-to-Market	-0.159^{***}	-0.096^{***}	-0.100^{***}	-0.352^{***}	-0.687^{***}
	(0.021)	(0.022)	(0.007)	(0.014)	(0.022)
Leverage	0.029^{**}	-0.106^{***}	-0.053^{***}	-0.281^{***}	-0.882^{***}
	(0.013)	(0.013)	(0.004)	(0.007)	(0.011)
Log Market Cap	0.385^{***}	-0.542^{***}	-0.400^{***}	-1.673^{***}	-3.892^{***}
	(0.050)	(0.059)	(0.017)	(0.035)	(0.055)
Momentum 12M	-0.154^{***}	-0.049^{***}	-0.035^{***}	-0.086^{***}	-0.109^{***}
	(0.006)	(0.006)	(0.002)	(0.003)	(0.005)
Turnover 60D	0.101^{***}	0.015^{*}	0.012^{***}	0.061^{***}	0.153^{***}
	(0.008)	(0.009)	(0.002)	(0.005)	(0.007)
Volatility 60D	-0.043^{***}	-0.011	-0.038^{***}	-0.061^{***}	-0.120^{***}
	(0.007)	(0.007)	(0.002)	(0.004)	(0.006)
Within R^2 (%)	2.07	1.48	5.96	16.65	33.67
Adj. R^2 (%)	1.97	1.38	5.86	16.57	33.60
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	stock	stock	stock	stock	stock

Notes: Standard errors clustered at stock level. All regressions include firm and month fixed effects. Accounting controls (ROA, operating profitability, cash/assets, sales growth, and log firm age) are not included because they are largely absorbed by firm fixed effects (limited within-firm variation) and can induce collinearity/unstable estimates. All independent variables are standardized (log-transformed variables not standardized). VIF check: All 5 PCA components have VIF ≤ 5. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: Video Content Characteristics and Abnormal Volatility (AV/CAV)

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	0.002	−0.028	0.0005	0.003***	0.008**
	(0.044)	(0.045)	(0.0005)	(0.001)	(0.003)
Title Sentiment Score	0.040	−0.156***	−0.0009***	0.0005	0.0006
	(0.035)	(0.034)	(0.0005)	(0.0009)	(0.003)
Text Sentiment Score	−0.029	−0.057	−0.0005	0.0007	0.002
	(0.034)	(0.036)	(0.0005)	(0.0009)	(0.002)
Text Richness Score	−0.146***	−0.017	−0.002***	−0.011***	−0.031***
	(0.035)	(0.036)	(0.0005)	(0.0009)	(0.002)
Influencer Reputation	0.038	0.086*	0.002***	0.010***	0.027***
	(0.044)	(0.045)	(0.0005)	(0.001)	(0.003)
Control Variables					
Book-to-Market	0.456*	−0.776***	−0.017***	−0.130***	−0.329***
	(0.252)	(0.257)	(0.003)	(0.007)	(0.019)
Leverage	−0.568***	−0.281*	0.006***	0.039***	0.145***
	(0.147)	(0.149)	(0.001)	(0.004)	(0.009)
Log Market Cap	−1.373**	−3.384***	−0.012*	−0.093***	−0.184***
	(0.582)	(0.642)	(0.006)	(0.017)	(0.044)
Momentum 12M	−0.598***	0.131*	−0.011***	−0.047***	−0.118***
	(0.070)	(0.072)	(0.0007)	(0.002)	(0.004)
Turnover 60D	−0.337***	−0.007	0.003***	0.028***	0.016**
	(0.098)	(0.099)	(0.0009)	(0.002)	(0.006)
Volatility 60D	1.486***	−0.872***	−0.022***	−0.092***	−0.189***
	(0.077)	(0.076)	(0.0007)	(0.002)	(0.005)
Within R^2 (%)	5.66	5.13	14.55	31.37	41.61
Adj. R^2 (%)	5.56	5.03	14.46	31.30	41.55
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	Stock	Stock	Stock	Stock	Stock

Notes: This table presents estimates from regression analyses of video content characteristics on stock volatility. Standard errors are clustered at the stock level. Analyses include firm, year, month fixed effects. Accounting controls (ROA, operating profitability, cash/assets, sales growth, and log firm age) are not included because they are largely absorbed by firm fixed effects (limited within-firm variation) and can induce collinearity/unstable estimates. All independent variables are standardized (log-transformed variables are not standardized). VIF check: All 5 PCA components have VIF ≤ 5. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Video Content Characteristics and Abnormal Volume (AVOL/CAVOL)

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	-0.035^{***}	0.023^{*}	0.002^{***}	0.001	0.006^{*}
	(0.012)	(0.013)	(0.0005)	(0.001)	(0.003)
Title Sentiment Score	0.038^{***}	0.015	0.0008^{**}	0.006^{***}	0.016^{***}
	(0.010)	(0.010)	(0.0005)	(0.0010)	(0.003)
Text Sentiment Score	-0.010	-0.022^{**}	0.0005	0.0005	-0.003
	(0.009)	(0.010)	(0.0005)	(0.0010)	(0.003)
Text Richness Score	0.065^{***}	0.052^{***}	0.0005	-0.002	0.002
	(0.010)	(0.010)	(0.0005)	(0.001)	(0.003)
Influencer Reputation	0.056^{***}	0.042^{***}	0.0005	0.008^{***}	0.016^{***}
	(0.011)	(0.012)	(0.0005)	(0.001)	(0.003)
Control Variables					
Book-to-Market	-0.232^{**}	-0.495^{***}	-0.013^{***}	-0.091^{***}	-0.298^{***}
	(0.114)	(0.122)	(0.004)	(0.015)	(0.040)
Leverage	-1.141^{***}	-1.026^{***}	-0.033^{***}	-0.137^{***}	-0.333^{***}
	(0.040)	(0.049)	(0.001)	(0.004)	(0.011)
Log Market Cap	-3.234^{***}	-2.856^{***}	-0.099^{***}	-0.515^{***}	-1.620^{***}
	(0.265)	(0.286)	(0.010)	(0.034)	(0.087)
Momentum 12M	0.281^{***}	0.135^{***}	0.006^{***}	0.034^{***}	0.121^{***}
	(0.022)	(0.024)	(0.0007)	(0.002)	(0.006)
Turnover 60D	0.042	-0.130^{***}	-0.009^{***}	-0.084^{***}	-0.489^{***}
	(0.034)	(0.036)	(0.001)	(0.004)	(0.009)
Volatility 60D	0.293^{***}	-0.127^{***}	-0.007^{***}	-0.037^{***}	-0.071^{***}
	(0.022)	(0.022)	(0.0007)	(0.002)	(0.006)
Within R^2 (%)	9.53	5.96	13.02	22.91	34.46
Adj. R^2 (%)	9.43	5.86	12.92	22.82	34.39
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	stock	stock	stock	stock	stock

Notes: Standard errors clustered at stock level. All regressions include firm and month fixed effects. Accounting controls (ROA, operating profitability, cash/assets, sales growth, and log firm age) are not included because they are largely absorbed by firm fixed effects (limited within-firm variation) and can induce collinearity/unstable estimates All independent variables are standardized (log-transformed variables not standardized).VIF check: All 5 PCA components have VIF ≤ 5. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Appendix

Table A-1: YouTube Event Sample Construction

Item	Count
Finance-focused YouTube channels	25
Raw videos collected	3,609
Videos retained after filters	1,924
Videos with valid transcripts	1,924
Videos with valid thumbnails	1,370
Raw YouTube–firm events	14,745
Merged and valid YouTube–firm events	12,492
Unique firms / PERMNOs	87
Unique trading days	677
Events during trading hours	4,780
Events after close / non-trading days assigned to next trading day	9,965
Mean YouTube video events per firm	169.48
Median YouTube video events per firm	10

Notes: This table reports the sample construction process for the YouTube event dataset. Videos were collected from finance-focused YouTube channels and matched to publicly traded firms. Events released outside trading hours and events occurring on non-trading days were assigned to the next available trading day. Thumbnail-based regressions use the subset of video–firm events for which thumbnail features are available; when multiple firm-events originate from the same video, the video-level thumbnail score is assigned to each associated firm-event.

Table A-2: Validation of Visual Measures

Panel A. Pearson Correlations with Video Engagement

Visual Measure	Log Views	Log Likes	Log Comments
Thumbnail Sentiment	0.277***	0.271***	0.263***
Color Saturation	0.264***	0.271***	0.269***
Contrast	0.232***	0.236***	0.212***
Face Emotion	0.236***	0.229***	0.205***
OCR Sentiment	0.218***	0.208***	0.221***

Panel B. Visual Component Scores Across Thumbnail Sentiment Tails

Component	Low Decile	High Decile	Difference
Color Saturation	0.048	0.176	0.128
Contrast	0.118	0.389	0.270
Face Emotion Intensity	0.652	0.945	0.292
OCR Sentiment	-0.568	1.362	1.931

Notes: Panel A reports Pearson correlations between Thumbnail Sentiment, its underlying visual components, and video engagement. Engagement is measured as log(1 + views), log(1 + likes), and log(1 + comments). Panel B compares mean visual component scores across the bottom and top deciles of the Thumbnail Sentiment distribution.
Pearson correlations only. Stars: * p<0.10, ** p<0.05, *** p<0.01.

Table A-3: Construction of YouTube-Derived Content Variables

Composite Variable	**PC1 (%)**	**Var.**	**Input Variables**
Thumbnail Sentiment Score	35.4		thumbnail_extreme_terms_count, thumbnail_all_caps_share, thumbnail_exclamation_density, thumbnail_text_red_share, thumbnail_text_green_share, thumbnail_text_yellow_share, thumbnail_face_count, thumbnail_face_area_ratio, thumbnail_emotion_confidence
Title Sentiment Score	31.2		title_total_words, all_caps_share, exclamation_density, extreme_terms_count, numeric_promises_count
Text Sentiment Score	67.0		ticker_avg_vader_score, ticker_avg_lm_score
Text Richness Score	34.1		video_length_minutes, video_readability, video_lexical_diversity, ticker_mention_count, ticker_words_ratio, ticker_duration_ratio, unique_tickers_count
Influencer Reputation	46.0		channel_subscribers_millions, channel_age, channel_activity

Notes: This table reports the construction of the YouTube-derived content variables used in the regression analysis. Each composite measure is constructed as the first principal component (PC1) of the listed input variables after standardizing inputs to zero mean and unit variance. The second column reports the percentage of variance explained by PC1. Regressions use standardized composite measures.

Table A-4: First Principal Component Loadings

Input Variable	Thumbnail	Title	Text Sent.	Text Rich.	Influencer
thumbnail_extreme_terms_count	0.393				
thumbnail_all_caps_share	0.380				
thumbnail_exclamation_density	0.084				
thumbnail_text_red_share	0.330				
thumbnail_text_green_share	0.213				
thumbnail_text_yellow_share	0.160				
thumbnail_face_count	0.418				
thumbnail_face_area_ratio	0.335				
thumbnail_emotion_confidence	0.476				
title_total_words		0.619			
all_caps_share		0.593			
exclamation_density		0.291			
extreme_terms_count		0.379			
numeric_promises_count		0.192			
ticker_avg_vader_score			0.707		
ticker_avg_lm_score			0.707		
video_length_minutes				0.553	
video_readability				0.018	
video_lexical_diversity				-0.543	
ticker_mention_count				0.317	
ticker_words_ratio				-0.113	
ticker_duration_ratio				-0.113	
unique_tickers_count				0.522	
channel_subscribers_millions					0.542
channel_age					0.619
channel_activity					0.568

Notes: This table reports first-principal-component (PC1) loadings for each composite measure. Input variables are standardized to zero mean and unit variance prior to PCA estimation. Larger absolute loadings indicate stronger contributions, while the sign reflects the direction of association with the principal component. The table facilitates interpretation of the latent dimensions captured by the PCA-based measures.

Table A-5: Selected Underlying Components of YouTube-Derived Content Measures

Variable	N	Mean	Std. Dev.	P25	Median	P75
Thumbnail Color Saturation	1,924	0.067	0.072	0.000	0.046	0.116
Thumbnail Contrast	1,924	0.151	0.170	0.000	0.098	0.250
Face Emotion / Expression Intensity	1,924	0.645	0.300	0.300	0.668	0.968
OCR Sentiment	1,924	0.000	1.140	-1.267	0.107	0.944
All-Caps / Exclamation Intensity	1,924	0.072	0.083	0.000	0.053	0.118
Ticker VADER Sentiment	1,924	0.215	0.290	0.015	0.177	0.355
Ticker LM Sentiment	1,924	-0.029	0.333	-0.106	0.000	0.048

Notes: This table reports summary statistics for selected underlying components used to construct the YouTube-derived content measures. Thumbnail Color Saturation is the mean of thumbnail red, green, and yellow text-color shares. Thumbnail Contrast is the difference between the maximum and minimum of the red, green, and yellow text-color shares. Face Emotion / Expression Intensity is the thumbnail emotion-detection confidence score. OCR Sentiment is the first principal component of thumbnail extreme terms, all-caps share, and exclamation density. All-Caps / Exclamation Intensity is the mean of title and thumbnail all-caps share and exclamation density. Transcript VADER Sentiment and Transcript LM Sentiment are ticker-specific transcript sentiment measures based on the VADER and Loughran–McDonald dictionaries, respectively.

Table A-6: Top 25 Financial YouTube Channels Discussing Individual Stocks (from Jan 2021 to Dec 2024)

No.	Channel	Approx. Subscribers (M)	Content Focus
1	Meet Kevin	2.0	Stock picks, earnings reactions, macro commentary
2	Andrei Jikh	2.2	Stock portfolio transparency, company reviews
3	Graham Stephan	4.8	Occasional stock analysis alongside finance content
4	Tom Nash	0.55	Individual stock deep dives, valuation
5	Financial Education (Jeremy Lefebvre)	0.75	Stock suggestions and analysis
6	Value Investing with Sven Carlin	0.35	Fundamental value stock research
7	Matt Kohrs	0.45	Post-earnings and trading events
8	Joseph Carlson Show	0.25	Dividend and growth stock portfolio
9	Everything Money	0.39	Stock valuation (DCF, PE analysis)
10	Patrick Boyle on Finance	0.62	Market and stock-specific education
11	The Plain Bagel	0.95	Academic stock valuation and risk concepts
12	WhiteBoard Finance (Marko)	1.0	Company fundamentals and investor behavior
13	Ricky Gutierrez	1.1	Swing trading and specific stock setups
14	Steven Dux	0.35	Trading individual equities (penny stocks)
15	ClayTrader	0.63	Technical analysis of individual stocks
16	Kenan Grace	0.45	Retail stock picks and market updates
17	Chris Sain	0.35	Stock alerts and weekly trade ideas
18	InvestAnswers	0.55	Individual stocks & crypto cross-asset commentary
19	Roensch Capital	0.22	Post-earnings reactions and sentiment
20	Financial Education 2 (Jeremy Lefebvre)	0.35	Follow-ups on stock ideas and market updates
21	Red Panda Investing	0.12	Tech and value stock analysis
22	Dividend Bull	0.10	Dividend stock picks
23	Daniel Pronk	0.10	Fundamental stock analysis
24	PPC Ian	0.08	Dividend and tech stock investing

Continued on next page

Table A-6 – continued from previous page

No.	Channel	Approx. Subscribers (M)	Content Focus
25	Stock Moe	0.72	Individual stock discussions and forecasts

Table A-7: Definitions of Event-Study Variables

Variable	Definition
$RR_{i,t}$	Raw return: $RR_{i,t} = R_{i,t}$.
$CRR_i(W)$	Cumulative raw return: $CRR_i(W) = \sum_{\tau \in W} R_{i,t+\tau}$.
$AR_{i,t}$	Abnormal return: $AR_{i,t} = R_{i,t} - \hat{R}_{i,t}$, where $\hat{R}_{i,t}$ is the FF5 expected return.
$CAR_i(W)$	Cumulative abnormal return: $CAR_i(W) = \sum_{\tau \in W} AR_{i,t+\tau}$.
$AVOL_{i,t}$	Abnormal log volume: $AVOL_{i,t} = \log(Vol_{i,t}) - \log(\bar{Vol}_{i,t})$.
$CAVOL_i(W)$	Cumulative abnormal volume: $CAVOL_i(W) = \sum_{\tau \in W} AVOL_{i,t+\tau}$.
$AV_{i,t}$	Abnormal volatility: $AV_{i,t} = r_{i,t}^2 - \bar{r^2}_{i,t}$.
$CAV_i(W)$	Cumulative abnormal volatility: $CAV_i(W) = \sum_{\tau \in W} AV_{i,t+\tau}$.

Table A-8: Event-Window Definitions and Economic Interpretation

Window	Economic Interpretation
$t = 0$	Contemporaneous market association; captures attention shocks and same-day trading correlation.
$t + 1$	Next-day adjustment due to overnight information processing and delayed viewing.
$t + 2$	Short-horizon drift from ongoing video engagement and algorithmic amplification.
$t + 3$	Continued diffusion of narrative-based information and attention spillover.
$[t{+}1, t{+}5]$	One-week cumulative effect; short-term assimilation, sentiment reinforcement, or reversal.
$[t{+}1, t{+}22]$	One-month horizon; tests the persistence of video-induced price pressure or correction.
$[t{+}1, t{+}66]$	Longer-run adjustment; evaluates whether initial reactions reflect mispricing or enduring information.
$[t{+}4, t{+}15]$	Medium-run window after immediate noise subsides; captures delayed disagreement or sentiment decay.
$[t{+}4, t{+}30]$	One-month adjusted horizon, isolating disagreement-based trading and volatility spillovers.
$t - 1$	Pre-event baseline; checks for anticipation or leakage before video release.
$[t{-}1, t{-}5]$	One-week pre-trend test; verifies absence of systematic run-up prior to the event.
$[t{-}1, t{-}22]$	One-month pre-trend window; ensures reactions are not driven by pre-existing momentum.

Table A-9: Summary Statistics of Dependent Variables, Independent Variables, and Control Variables

Variable	N	Mean	Std. Dev.	P25	Median	P75
Panel A. Dependent Variables						
RR [t=0]	12,492	0.000	0.016	-0.009	0.000	0.009
RR [t+1]	12,492	0.000	0.016	-0.009	0.001	0.010
CRR [t+1, t+5]	12,492	0.002	0.037	-0.019	0.003	0.024
CRR [t+1, t+22]	12,492	0.009	0.081	-0.040	0.010	0.060
CRR [t+1, t+66]	12,492	0.033	0.131	-0.052	0.032	0.112
CAR [t-65, t-61]	12,492	-0.000	0.033	-0.018	-0.000	0.019
CAR [t-65, t-44]	12,492	-0.002	0.072	-0.044	-0.001	0.042
CAR [t-65, t-1]	12,492	-0.009	0.119	-0.087	-0.009	0.073
CAR [t-65, t=0]	12,492	-0.010	0.120	-0.089	-0.009	0.074
AR [t=0]	12,492	-0.001	0.014	-0.008	-0.001	0.007
AR [t+1]	12,492	-0.000	0.014	-0.008	-0.000	0.008
CAR [t+1, t+5]	12,492	-0.002	0.033	-0.019	-0.001	0.017
CAR [t+1, t+22]	12,492	-0.007	0.074	-0.049	-0.004	0.039
CAR [t+1, t+66]	12,492	-0.013	0.118	-0.094	-0.013	0.068
CAV [t-65, t-61]	12,492	0.0000	0.0257	-0.0161	-0.0025	0.0134
CAV [t-65, t-44]	12,492	-0.0011	0.0588	-0.0342	0.0000	0.0345
CAV [t-65, t-1]	12,492	0.0051	0.2127	-0.0884	0.0037	0.0965
CAV [t-65, t=0]	12,492	0.0033	0.2080	-0.0923	0.0041	0.0977
AV [t=0]	12,492	-0.0001	0.0141	-0.0073	-0.0022	0.0048
AV [t+1]	12,492	0.0005	0.0144	-0.0072	-0.0019	0.0053
CAV [t+1, t+5]	12,492	-0.0003	0.0386	-0.0200	-0.0024	0.0167
CAV [t+1, t+22]	12,492	0.0180	0.2314	-0.0577	0.0002	0.0545
CAV [t+1, t+66]	12,492	0.0137	0.4839	-0.1676	-0.0059	0.1633
CAVOL [t-65, t-61]	12,492	-0.00007	0.00875	-0.00513	-0.00102	0.00374
CAVOL [t-65, t-44]	12,492	-0.00142	0.02598	-0.01418	-0.00087	0.01165
CAVOL [t-65, t-1]	12,492	-0.01244	0.17096	-0.04465	-0.00271	0.03451
CAVOL [t-65, t=0]	12,492	-0.01253	0.17405	-0.04616	-0.00269	0.03553
AVOL [t=0]	12,492	-0.00015	0.00599	-0.00189	-0.00044	0.00116
AVOL [t+1]	12,492	0.00017	0.00563	-0.00188	-0.00049	0.00118
CAVOL [t+1, t+5]	12,492	0.00321	0.03840	-0.00790	-0.00135	0.00591
CAVOL [t+1, t+22]	12,492	0.04504	0.35754	-0.02754	-0.00197	0.02243
CAVOL [t+1, t+66]	12,492	0.00036	0.22308	-0.08276	-0.00348	0.06796
Panel B. Independent Variables						
Thumbnail Sentiment Score	12,492	-0.000	1.785	-2.508	0.591	1.346
Title Sentiment Score	12,492	0.000	1.249	-1.424	0.062	0.688
Text Sentiment Score	12,492	-0.000	1.158	-0.620	-0.095	0.491
Text Richness Score	12,492	0.000	1.545	-0.800	-0.101	0.928
Influencer Reputation	12,492	-0.000	1.174	-0.633	-0.633	0.437
Panel C. Control Variables						
Book-to-Market	12,492	0.409	0.254	0.178	0.411	0.615
Leverage	12,492	0.298	0.131	0.179	0.312	0.413
Log Market Cap	12,492	11.347	0.817	10.561	11.280	11.858
Momentum 12M	12,492	0.143	0.267	-0.035	0.100	0.265
Turnover 60D	12,492	0.006	0.004	0.004	0.006	0.007
Volatility 60D	12,492	0.016	0.005	0.012	0.015	0.019

Notes: This table reports summary statistics for the dependent variables, PCA-based YouTube content measures, and firm-level control variables used in the regression analysis. The sample consists of 12,492 unique YouTube video × ticker events from 2021–2024. PCA-based variables are the first principal components estimated from the video × ticker sample, are not winsorized, and have means approximately equal to zero by construction. All other variables are winsorized at the 1st and 99th percentiles. RR and CRR denote raw return and cumulative raw return, respectively; AR and CAR denote abnormal returns and cumulative abnormal returns, respectively; AV and CAV denote abnormal trading volume and cumulative abnormal trading volume, respectively; and AVOL and CAVOL denote abnormal volatility and cumulative abnormal volatility, respectively. All market outcome variables are reported in decimal form (0.01 = 1% for returns), consistent with CRSP returns. Summary statistics are reported for the unstandardized variables. Influencer Reputation is measured at the channel level.

Table A-10: Robustness Test for Video Content Characteristics and Abnormal Returns (AR/CAR) -Alternative Event Windows

Variable	[t = 0]	[t + 1]	[t + 2]	[t + 3]	[t + 4, t + 15]	[t + 4, t + 30]
Thumbnail Sentiment Score	-0.012^{***}	-0.008^{*}	-0.002	0.007^{*}	-0.021	0.012
	(0.004)	(0.004)	(0.004)	(0.004)	(0.016)	(0.023)
Title Sentiment Score	-0.001	-0.010^{***}	-0.004	0.008^{**}	-0.038^{***}	-0.020
	(0.003)	(0.003)	(0.003)	(0.003)	(0.013)	(0.017)
Text Sentiment Score	0.012^{***}	0.003	0.014^{***}	-0.007^{**}	-0.005	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.013)	(0.019)
Text Richness Score	-0.004	0.008^{**}	-0.011^{***}	-0.007^{**}	-0.011	0.060^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)	(0.020)
Influencer Reputation	0.010^{***}	-0.003	0.004	-0.002	-0.025	-0.080^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.016)	(0.023)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	2.07	1.48	2.22	2.41	10.42	17.75
Adj. R^2 (%)	1.97	1.38	2.12	2.30	10.32	17.66
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	Stock	Stock	Stock	Stock	Stock	Stock

Notes: This table presents regression results of video content characteristics on abnormal returns (AR) and cumulative abnormal returns (CAR) across different event windows. Abnormal returns are computed as the difference between realized returns and expected returns based on the Fama–French five-factor model. Cumulative abnormal returns are aggregated over the specified event windows. Standard errors are clustered at the stock level. All regressions include firm and calendar-time fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, Momentum (12M), Turnover (60D), and Volatility (60D). All independent variables are standardized (log-transformed variables are not standardized). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A-11: Robustness Test for Video Content Characteristics and Abnormal Volatility (AV/CAV)-Alternative Event Windows

Variable	[t = 0]	[t + 1]	[t + 2]	[t + 3]	[t + 4, t + 15]	[t + 4, t + 30]
Thumbnail Sentiment Score	0.002	-0.028	0.189^{***}	-0.016	0.070^{***}	0.026
	(0.044)	(0.045)	(0.045)	(0.044)	(0.021)	(0.036)
Title Sentiment Score	0.040	-0.156^{***}	-0.109^{***}	-0.070^{**}	0.009	-0.0005
	(0.035)	(0.034)	(0.035)	(0.034)	(0.017)	(0.029)
Text Sentiment Score	-0.029	-0.057	0.054	0.037	0.018	0.032
	(0.034)	(0.036)	(0.036)	(0.038)	(0.016)	(0.028)
Text Richness Score	-0.146^{***}	-0.017	-0.048	-0.121^{***}	-0.202^{***}	-0.303^{***}
	(0.035)	(0.036)	(0.037)	(0.036)	(0.016)	(0.028)
Influencer Reputation	0.038	0.086^{*}	0.009	0.058	0.119^{***}	0.381^{***}
	(0.044)	(0.045)	(0.043)	(0.043)	(0.020)	(0.035)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	5.66	5.13	5.90	4.52	24.23	34.23
Adj. R^2 (%)	5.56	5.03	5.80	4.42	24.15	34.16
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	Stock	Stock	Stock	Stock	Stock	Stock

Notes: This table reports regression results of video content characteristics on abnormal volatility measures. Abnormal volatility (AV) is defined using squared abnormal returns, and cumulative abnormal volatility (CAV) is constructed over the specified event windows. Standard errors (in parentheses) are clustered at the stock level. All regressions include firm and calendar-time fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, Momentum (12M), Turnover (60D), and Volatility (60D). All independent variables are standardized (log-transformed variables are not standardized). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A-12: Robustness Test Result for Video Content Characteristics and Abnormal Volume (AVOL/CAVOL) -Alternative Event Windows

Variable	[t = 0]	[t + 1]	[t + 2]	[t + 3]	[t + 4, t + 15]	[t + 4, t + 30]
Thumbnail Sentiment Score	-0.035^{***}	0.023^{*}	0.267^{***}	0.257^{***}	-0.457^{***}	-0.919^{***}
	(0.012)	(0.013)	(0.040)	(0.048)	(0.057)	(0.107)
Title Sentiment Score	0.038^{***}	0.015	-0.061^{***}	-0.091^{***}	0.190^{***}	0.355^{***}
	(0.010)	(0.010)	(0.011)	(0.024)	(0.023)	(0.047)
Text Sentiment Score	-0.010	-0.022^{**}	0.007	-0.054^{*}	0.039	0.064
	(0.009)	(0.010)	(0.018)	(0.032)	(0.041)	(0.067)
Text Richness Score	0.065^{***}	0.052^{***}	0.019	-0.040	0.032	-0.026
	(0.010)	(0.010)	(0.029)	(0.031)	(0.042)	(0.106)
Influencer Reputation	0.056^{***}	0.042^{***}	-0.191^{***}	-0.195^{***}	0.437^{***}	0.889^{***}
	(0.011)	(0.012)	(0.031)	(0.036)	(0.043)	(0.078)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	9.53	5.96	1.86	1.67	3.38	3.92
Adj. R^2 (%)	9.43	5.86	1.75	1.57	3.27	3.81
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
Cluster	Stock	Stock	Stock	Stock	Stock	Stock

Notes: This table presents regression results of video content characteristics on abnormal trading activity. Abnormal volume (AVOL) is measured using abnormal log trading volume, and cumulative abnormal volume (CAVOL) aggregates deviations over the specified event windows. Standard errors are clustered at the stock level. All regressions include firm and calendar-time fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, Momentum (12M), Turnover (60D), and Volatility (60D). All independent variables are standardized (log-transformed variables are not standardized). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A-13: Robustness Test Result Video For Content Characteristics and Abnormal Returns (AR/CAR) — PSM

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	-0.0096	-0.0104	-0.0039^{*}	-0.0073	-0.0061
	(0.0085)	(0.0085)	(0.0022)	(0.0045)	(0.0075)
Title Sentiment Score	0.0095	0.0080	0.0051^{***}	0.0036	0.0217^{***}
	(0.0088)	(0.0069)	(0.0019)	(0.0042)	(0.0073)
Text Sentiment Score	0.0053	0.0046	0.0009	0.0034	0.0059
	(0.0065)	(0.0065)	(0.0018)	(0.0035)	(0.0059)
Text Richness Score	-0.0038	0.0039	-0.0019	-0.0069^{*}	-0.0100^{*}
	(0.0067)	(0.0068)	(0.0018)	(0.0035)	(0.0058)
Influencer Reputation	0.0093	0.0007	0.0009	-0.0018	-0.0180^{**}
	(0.0089)	(0.0086)	(0.0022)	(0.0047)	(0.0079)
Control Variables	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	1.2636	1.2471	3.7213	11.8859	24.6007
Adj. R^2 (%)	0.7612	0.7446	3.2315	11.4376	24.2171
Fixed Effects	firm, month, year	firm, month, year	firm, month, year	firm, month, year	firm, month, year
Cluster	Stock	Stock	Stock	Stock	Stock

Notes: This table presents estimates from regression analyses of video content characteristics on stock returns. Standard errors (in parentheses) are clustered at the Matched pair and stock level. Analyses include firm, month, year fixed effects. Standard firm-level controls are included in all specifications; see text for variable definitions. All independent variables are standardized (log-transformed variables are not standardized). Thumbnail characteristics include all caps share, text color shares, and exclamation density. Content characteristics include numeric promises count, extreme terms count, total words, and all caps share. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-14: Video Content Characteristics and Abnormal Volatility (AV/CAV) — PSM

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	-0.2147^{**}	-0.0610	$-5.6587e-05$	-0.0005	-0.0011
	(0.1015)	(0.1012)	(0.0009)	(0.0024)	(0.0069)
Title Sentiment Score	0.1470	-0.1510^{*}	$3.0528e-05$	0.0002	-0.0103^{**}
	(0.1073)	(0.0828)	(0.0007)	(0.0019)	(0.0050)
Text sentiment Score	-0.0561	-0.0549	0.0005	0.0011	0.0035
	(0.0781)	(0.0793)	(0.0007)	(0.0019)	(0.0051)
Text Richness Score	-0.0083	-0.0288	0.0004	0.0014	0.0015
	(0.0784)	(0.0790)	(0.0007)	(0.0018)	(0.0050)
Influencer Reputation	0.1655	0.1353	0.0018^{**}	0.0065^{***}	0.0181^{***}
	(0.1060)	(0.1028)	(0.0009)	(0.0025)	(0.0070)
Control Variables	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	4.1685	4.3273	14.4921	30.8106	41.5002
Adj. R^2 (%)	3.6810	3.8405	14.0571	30.4586	41.2026
Fixed Effects	firm, month, year	firm, month, year	firm, month, year	firm, month, year	firm, month, year
Cluster	Stock	Stock	Stock	Stock	Stock

Notes: This table presents estimates from regression analyses of video content characteristics on stock volatility. Standard errors (in parentheses) are clustered at the Matched pair and stock level. Analyses include firm, month, year fixed effects. Standard firm-level controls are included in all specifications; see text for variable definitions. All independent variables are standardized (log-transformed variables are not standardized). Thumbnail characteristics include all caps share, text color shares, and exclamation density. Content characteristics include numeric promises count, extreme terms count, total words, and all caps share.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-15: Video Content Characteristics and Abnormal Volume (AVOL/CAVOL) — PSM

Variable	[t = 0]	[t + 1]	[t + 1, t + 5]	[t + 1, t + 22]	[t + 1, t + 66]
Thumbnail Sentiment Score	-0.0317	-0.0006	0.0004	0.0002	0.0105
	(0.0295)	(0.0297)	(0.0008)	(0.0027)	(0.0073)
Title Sentiment Score	0.1010^{***}	-0.0030	0.0001	0.0020	0.0042
	(0.0337)	(0.0224)	(0.0007)	(0.0022)	(0.0060)
Text sentiment Score	0.0039	-0.0210	0.0004	0.0005	0.0007
	(0.0235)	(0.0228)	(0.0007)	(0.0022)	(0.0061)
Text Richness Score	0.0293	0.0298	0.0008	0.0022	0.0033
	(0.0237)	(0.0242)	(0.0006)	(0.0021)	(0.0054)
Influencer Reputation	0.0727^{**}	0.0743^{**}	0.0005	0.0048^{*}	0.0075
	(0.0305)	(0.0304)	(0.0008)	(0.0026)	(0.0069)
Control Variables	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	5.5626	5.3326	9.1903	17.1656	25.8243
Adj. R^2 (%)	5.0821	4.8510	8.7283	16.7442	25.4469
Fixed Effects	firm, month, year	firm, month, year	firm, month, year	firm, month, year	firm, month, year
Cluster	Stock	Stock	Stock	Stock	Stock

Notes: This table presents estimates from regression analyses of video content characteristics on stock volume. Standard errors (in parentheses) are clustered at the Matched pair and stock level. Analyses include firm, month, year fixed effects. Standard firm-level controls are included in all specifications; see text for variable definitions. All independent variables are standardized (log-transformed variables are not standardized). Thumbnail characteristics include all The large negative coefficient on Log Market Cap in abnormal volume regressions reflects the mechanical relationship whereby large-cap stocks have higher absolute trading volumes but lower proportional volume volatility. This does not diminish the economic significance of YouTube content variables, which capture incremental information effects beyond firm size characteristics. caps share, text color shares, and exclamation density. Content characteristics include numeric promises count, extreme terms count, total words, and all caps share.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-16: Placebo Test for Video Content Characteristics and Abnormal Returns (AR/CAR)

Variable	$[t-66, t-53]$	$[t-52, t-39]$	$[t-38, t-25]$	$[t-24, t-11]$	$[t-10, t-1]$
Thumbnail Sentiment Score	0.000	-0.000	-0.000	0.002*	0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Title Sentiment Score	0.000	0.000	0.000**	-0.001	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Text Sentiment Score	-0.001	0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Text Richness Score	-0.001	0.000	-0.000	0.000	0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Influencer Reputation	0.002	-0.000	-0.001**	-0.002	-0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Control Variables	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	5.28	7.39	8.91	8.69	8.42
Adj. R^2 (%)	4.06	6.21	7.74	7.52	7.24
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
firm, year, month					
Cluster	Stock	Stock	Stock	Stock	Stock
Stock					

Notes: This table reports placebo-test regressions of video content characteristics on abnormal stock returns using pre-event windows. Standard errors (in parentheses) are clustered at the stock level. All specifications include firm and time (year and month) fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, 12-month Momentum, 60-day Turnover, and 60-day Volatility. All independent variables are standardized (log-transformed variables are not standardized). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-17: Placebo Test Result for Video Content Characteristics and Abnormal Volatility (AV/CAV)

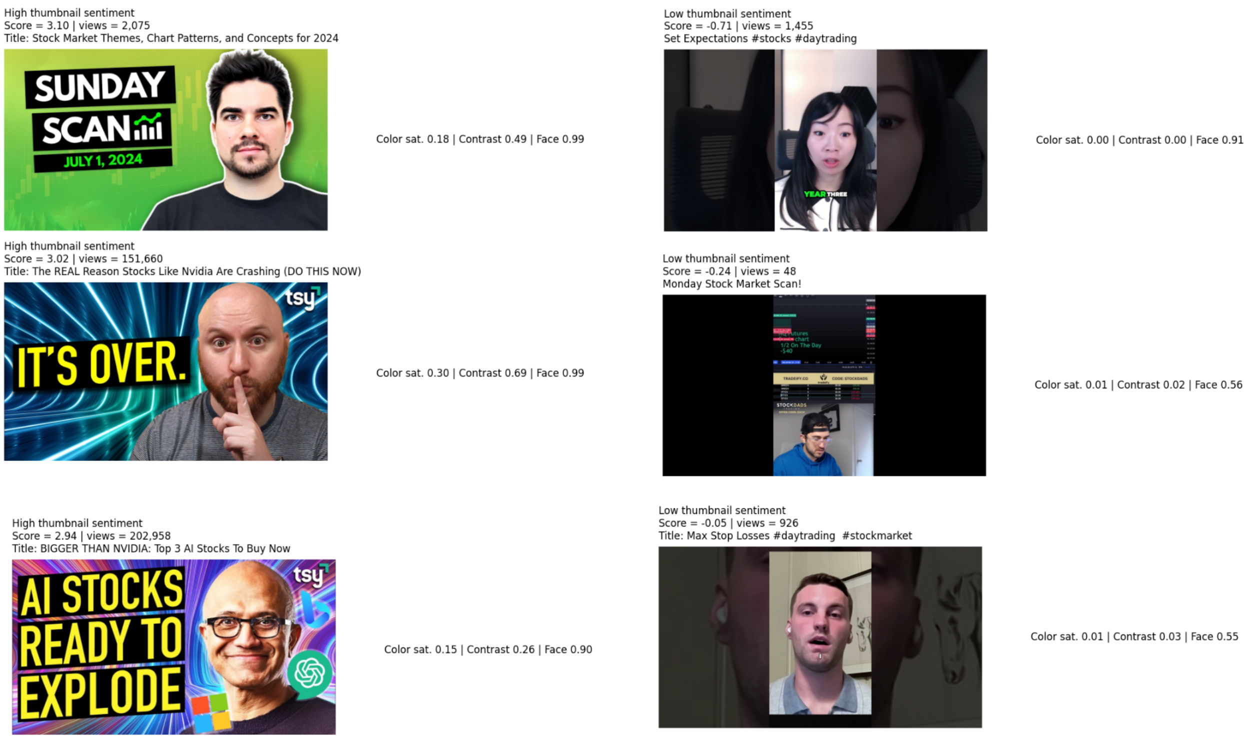
Variable	$[t-66, t-53]$	$[t-52, t-39]$	$[t-38, t-25]$	$[t-24, t-11]$	$[t-10, t-1]$
Thumbnail Sentiment Score	0.000	0.001	0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Title Sentiment Score	-0.001	-0.001	-0.001	-0.001	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Text Sentiment Score	-0.001	-0.000	0.001*	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Text Richness Score	-0.001	-0.000	-0.001	0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Influencer Reputation	-0.002**	-0.001	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	34.67	49.77	41.71	46.57	45.42
Adj. R^2 (%)	33.83	49.13	40.97	45.89	44.72
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
firm, year, month					
Cluster	Stock	Stock	Stock	Stock	Stock
Stock					

Notes: This table reports placebo-test regressions of video content characteristics on abnormal stock volatility using pre-event windows. Standard errors (in parentheses) are clustered at the stock level. All specifications include firm and time (year and month) fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, 12-month Momentum, 60-day Turnover, and 60-day Volatility. All independent variables are standardized (log-transformed variables are not standardized).
Placebo design: regressions are estimated over pre-event windows to test for pre-trends. The absence of consistent statistically significant coefficients suggests no systematic relationship prior to the event, supporting the identification strategy.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-18: Placebo Tests for Video Content Characteristics and Abnormal Volume (AVOL/CAVOL)

Variable	$[t-66, t-53]$	$[t-52, t-39]$	$[t-38, t-25]$	$[t-24, t-11]$	$[t-10, t-1]$
Thumbnail Sentiment Score	0.000	0.000	0.001**	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Title Sentiment Score	0.001*	0.000	0.000	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Text Sentiment Score	0.000	0.000	0.000	0.000	-0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Text Richness Score	0.001	0.000	0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Influencer Reputation	-0.001	0.000	-0.001	0.001	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes	Yes	Yes
Within R^2 (%)	14.84	19.37	14.52	17.33	19.76
Adj. R^2 (%)	13.74	18.34	13.43	16.27	18.73
Fixed Effects	firm, year, month	firm, year, month	firm, year, month	firm, year, month	firm, year, month
firm, year, month					
Cluster	Stock	Stock	Stock	Stock	Stock
Stock					

Notes: This table reports placebo-test regressions of video content characteristics on abnormal trading volume using pre-event windows. Standard errors (in parentheses) are clustered at the stock level. All specifications include firm and time (year and month) fixed effects. Control variables include Book-to-Market, Leverage, Log Market Capitalization, 12-month Momentum, 60-day Turnover, and 60-day Volatility. All independent variables are standardized (log-transformed variables are not standardized).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.



(a) High Thumbnail Sentiment

(b) Low Thumbnail Sentiment

Figure A-1: Thumbnail Sentiment Examples

Notes: Panel (a) presents illustrative examples of videos with high Thumbnail Sentiment scores. These thumbnails are more vivid, with stronger color contrast and facial cues. Panel (b) presents illustrative examples of videos with low Thumbnail Sentiment scores. These thumbnails are visually plainer, with lower color saturation and contrast.