Finfluencers and Stock Market Quality: Evidence from TikTok

Qingfu Liu Yang Song Chuanjie Wang Jiaer Yang*

May 31, 2026

Abstract

We provide first causal evidence on how financial influencers (finfluencers) affect stock market quality. Exploiting a nationwide crackdown on misleading online financial content in China, we study how restricting finfluencer activity on China's TikTok affects stock market quality. Using the platform's TikTok Index to measure stock-level exposure to finfluencer content, we implement a difference-in-differences design. We find that stocks with higher pre-campaign exposure experience a significant decline in intraday volatility after the crackdown. Mechanism analyses suggest that the policy improves the retail information environment and dampens attention-driven retail participation. The effects are stronger where formal information channels are weaker and among stocks with more male-dominated audiences. These findings show that creator-driven financial information shapes retail trading and stock market quality.

Keywords: Finfluencers; Social Media, Market Quality, Retail Investors.

JEL Classification: G14, G18, D83.

*Qingfu Liu is at the School of Economics, Fudan University and the Shanghai Institute for Mathematics and Interdisciplinary Sciences; Email: liuqf@fudan.edu.cn. Yang Song is at the Foster School of Business, University of Washington; Email: songy18@uw.edu. Chuanjie Wang is at the School of Finance, Shanghai University of Finance and Economics; Email: wangchuanjie@mail.shufe.edu.cn. Jiaer Yang is at the School of Economics, Fudan University; Email: jeyang24@m.fudan.edu.cn.

1 Introduction

Financial influencers (**Finfluencers**) are reshaping how retail investors acquire market information by packaging investment views for large online audiences (Merkley et al., 2024; Hull and Qi, 2024). However, the market consequences of this influence remain theoretically ambiguous. Finfluencers may broaden access to timely market information, but they may also amplify financial misinformation and expose retail investors to manipulation (Kogan et al., 2023; Merkley et al., 2024). Whether finfluencers ultimately improve price efficiency or impair stock market quality is therefore a central empirical question.

Traditional market participants rely on institutional channels to disseminate financial information (Engelberg and Parsons, 2011; Blankespoor et al., 2014). Finfluencers operate as a distinct intermediary governed by algorithmic platforms. Since platform architecture fundamentally shapes the resulting social signal (Cookson et al., 2024), this structural shift in information supply yields competing theoretical predictions. These finfluencers might improve market efficiency by diffusing value-relevant content and translating complex public signals (Antweiler and Frank, 2004; Chen et al., 2014). Alternatively, engagement-driven environments frequently prioritize sensational narratives. Such mechanisms can propagate manipulated information and magnify market stress by sustaining segmented beliefs (Cookson et al., 2023; Kogan et al., 2023; Cookson et al., 2026).

A definitive answer requires causal evidence on the transmission from finfluencer content to stock market quality. Extensive research establishes that social media platforms shape investor attention and belief formation (Antweiler and Frank, 2004; Chen et al., 2014; Cookson and Niessner, 2020). However, isolating the market effects of creator-driven information supply remains empirically difficult. Creator-generated content is inherently endogenous. Finfluencers naturally target volatile or newsworthy stocks, which confounds platform exposure with reverse causality. Furthermore, standard empirical proxies contain substantial measurement error because they capture broad investor demand rather than

specific creator supply (deHaan et al., 2025). As a result, standard observational data cannot cleanly isolate finfluencer information shocks from general stock salience.

To overcome this empirical challenge, we exploit a nationwide regulatory intervention in China. On August 27, 2021, the Cyberspace Administration of China (CAC) launched the Clean and Orderly Cyberspace campaign to rectify the unauthorized dissemination and distortion of financial information. Although the mandate covered both commercial websites and self-media platforms, our textual evidence shows that targeted financial-advice language remained persistently low in news outlets and commercial websites. This low baseline points to platform-based creators as the main binding constraint of the policy. We therefore focus on China's TikTok (Douyin), where finfluencers deliver retail-facing stock advice through push-based algorithmic feeds. The crackdown generated a sharp contraction in this specific source of information supply, which allows us to identify the causal effect of finfluencer regulation on stock market quality.

Our empirical analysis covers A-share firms listed on the Shanghai and Shenzhen main boards. We use an event window from February 27, 2021 to March 26, 2022 around the August 27 intervention, yielding 634,902 stock-day observations for 2,501 stocks. We quantify stock-level visibility on China's TikTok using Ocean Engine's TikTok Index. Unlike search-based measures, which mainly capture active investor demand, this feed-based index is better suited to capturing algorithmic amplification and creator-driven supply. We define policy exposure based on each stock's average TikTok Index during the pre-intervention month and classify stocks above the cross-sectional median as the high-exposure treatment group. To evaluate market quality, we construct realized volatility and realized range volatility from five-minute intraday prices. These measures capture short-horizon price instability (O'Hara and Ye, 2011; Malinova and Park, 2015) and information-related trading frictions in equity markets (Comerton-Forde et al., 2019; Eaton et al., 2022).

The difference-in-differences estimates document a pronounced causal effect of the crackdown. After the campaign, stocks with high pre-campaign finfluencer exposure experience

a significant decline in intraday volatility. Realized volatility and realized range volatility fall by 13.4% and 10.5% of their sample means, respectively. Event-study estimates show no differential pre-trends and indicate a sustained improvement in short-horizon price stability. To isolate this creator-driven shock from confounding structures, we address observable covariate imbalances and functional-form dependence through entropy balancing and double debiased machine learning. Sensitivity analyses further ensure that unobserved confounders cannot overturn this volatility reduction. Our design also separates creator-driven information supply from general stock salience. The results are not replicated by placebo exposure based on conventional newspaper coverage and remain robust after residualizing platform exposure against formal information demand. The evidence indicates that regulating finfluencers improves stock market quality by reducing volatility.

To explain why regulating finfluencers improves market quality, we first examine an information-quality channel rooted in content supply. Creator-level data reveal a clear post-intervention shift away from stock-specific commentary toward broader industry narratives. By curtailing speculative promotion, this supply-side pivot changes the retail information environment. High-exposure stocks subsequently experience substantial increases in engagement-weighted post quality and stronger sentiment conformity. This convergence indicates a sharp reduction in retail disagreement, a mechanism that can dampen noise-driven trading and reduce non-fundamental price volatility (Antweiler and Frank, 2004; Cookson and Niessner, 2020). The evidence suggests that the policy improves market quality by replacing speculative noise with a more coherent information environment.

A second channel operates through investor attention and participation. Platform-level metrics show a sharp contraction in retail engagement after the intervention, with significant declines in both algorithmic visibility and active search popularity on TikTok. This muted platform exposure directly diminishes subsequent retail activity, driving a substantial decrease in total post volume alongside a lower small-order trading volume. Because retail trading is strongly shaped by attention constraints (Barber and Odean,

2000; Da et al., 2011), this decline in platform visibility limits the entry of attention-driven traders. Reduced retail participation can thus stabilize intraday prices (Foucault et al., 2011; Eaton et al., 2022). This participation channel works alongside the improvement in information quality to explain the aggregate gain in stock market quality.

The cross-sectional patterns are consistent with these two mechanisms. The volatility decline is significantly stronger among firms with weak formal information environments. When professional intermediary coverage is limited or disclosure quality is low, traditional information channels provide less discipline over investor beliefs and exacerbate information frictions (Leuz and Verrecchia, 2000; Lang et al., 2003). Retail investors in these settings are more likely to rely on informal social signals, making prices more exposed to finfluencer-driven noise. Opaque information environments also leave stock prices more susceptible to non-fundamental volatility (Durnev et al., 2003; Jin and Myers, 2006). Removing platform-amplified noise should therefore generate the largest stabilizing benefits where formal information is weakest. Audience demographics provide further support for the attention-driven channel. The market improvement is markedly stronger for stocks with predominantly male finfluencer audiences. Because male retail investors trade more aggressively and respond more strongly to salient external stimuli (Barber and Odean, 2001; Arnold et al., 2022), the stabilizing effect of the intervention is concentrated where platform content is most likely to translate into speculative trading.

Our paper makes three contributions. First, we provide the first causal evidence on the market-quality consequences of finfluencers. Digital influencers operate as powerful information intermediaries that shape user engagement and economic choices (Mallipeddi et al., 2022; Nistor and Selove, 2024). In financial markets, finfluencers use persuasive narratives to influence retail trading behavior (Merkley et al., 2024; Hull and Qi, 2024). Their incentives can favor audience growth and platform monetization over informational accuracy (Kakhbod et al., 2023; Nistor et al., 2025). We extend this individual-level behavioral perspective to market-level price formation. Using a regulatory intervention that

constrained unregulated platform commentary, we show that heavily exposed stocks experience a substantial decline in intraday volatility. This evidence moves the emerging influencer literature beyond retail responses by identifying market-wide consequences for price stability and market quality.

Second, we broaden the social media finance literature by studying feed-based short video networks. Prior finance and accounting research has largely examined text-centric forums, while platform economics and marketing research show that creator-driven ecosystems follow different organizing principles. Strategic influencer scheduling and monetization incentives shape content supply (Mallipeddi et al., 2022; Nistor et al., 2025; Liang et al., 2025). Algorithmic recommendation systems and engagement metrics then determine which content reaches users at scale (Bhargava, 2022; Qian and Jain, 2024; Yang et al., 2024, 2025). We bring this platform architecture into financial market analysis. TikTok does not simply reflect investor demand. It supplies algorithmically curated stock narratives to retail audiences. Regulating this information supply reduces speculative retail participation and stabilizes intraday prices. This regulatory shock identifies short video platforms as a distinct information infrastructure in modern retail finance.

Third, we contribute to the market microstructure literature by identifying an upstream source of retail-driven volatility. Traditional research studies market quality largely through exchange architecture and trading design (O'Hara and Ye, 2011). Related work shows that trading rules and liquidity provision shape pricing efficiency (Malinova and Park, 2015; Comerton-Forde et al., 2019). In this framework, retail participation is often treated as an exogenous source of attention-driven noise (Barber and Odean, 2008). Such uninformed flow can consume liquidity and destabilize intraday prices (Eaton et al., 2022). We show that retail-driven volatility is partly organized upstream by platforms that curate investor attention. Regulating algorithmic financial content reduces volatility for heavily exposed stocks, placing digital platforms alongside trading venues as key determinants of high-frequency price stability.

The rest of the paper proceeds as follows. Section 2 describes the institutional background and the CAC Finfluencer Campaign. Section 3 introduces the data and variable construction. Section 4 presents the main empirical results and identification tests. Section 5 examines the information-quality and retail-attention mechanisms. Section 6 reports heterogeneity analyses and robustness tests. Section 7 concludes.

2 Institutional Background and Regulatory Setting

Financial influencers (finfluencers) operate as powerful digital intermediaries that shape retail information exposure (Merkley et al., 2024; Hull and Qi, 2024). Unlike traditional financial media, finfluencers rely on creator-based production, platform recommendation, and audience engagement to distribute market views to broad retail audiences. Their growing influence raises concerns that unlicensed or weakly supervised financial commentary can expose investors to low-quality advice, amplify speculative narratives, and generate negative externalities for market quality.[1] We know little about the market-level economic effects of finfluencer activity.

On August 27, 2021, the Cyberspace Administration of China launched the *Clean and Orderly Cyberspace: Special Rectification of Financial Information on Commercial Websites and Self-Media* campaign. The campaign targeted misleading, manipulative, sensational, and unlicensed financial content disseminated through commercial websites and self-media accounts.[2] It compelled digital platforms to rigorously verify information sources and immediately dismantle non-compliant accounts. Consequently, by internalizing com-

[1] See International Organization of Securities Commissions (IOSCO, 2025), *Finfluencers* (FR/08/2025), and World Economic Forum (2024), "Are 'finfluencers' the future of financial advice?". Recent media and practitioner discussions also highlight the risks associated with financial advice on social media, including misleading or promotional content and potential regulatory exposure. See *The New York Times* (2025), "Financial Advice on Social Media Is Growing, and Risky", and *Harvard Business Review* (2025), "How Finfluencers Can Create Risk for Your Company".

[2] The campaign identified eight categories of problematic financial content, including distorted interpretations of economic policies, narratives undermining China's economic or financial outlook, rumor fabrication and amplification, manipulative rewrites of otherwise compliant financial news, unlicensed stock recommendations and price manipulation, improper use of negative information for private gain, emotion-driven sensationalism, and impersonation of officials or experts.

pliance costs at the platform level, the campaign generated a sharp contraction in the supply of unregulated financial advice.[3]

The campaign was broad in legal scope and covered both commercial financial websites and self-media accounts. These two channels differ in how financial content is produced and screened. Commercial financial websites or information platforms typically operate under routine baseline review protocols. Self-media accounts instead rely on decentralized creator production and largely depend on creators' own compliance discipline. This distinction suggests that the policy should be more binding for self-media creators than for commercial financial websites.

[INSERT FIGURE 1 HERE]

To assess this conjecture, we analyze 176,269 articles published by 744 nonofficial news providers and commercial financial websites in the CSMAR database over the 12 months surrounding the intervention. We deploy three leading Chinese large language models (DeepSeek V4 Flash, ERNIE 4.5 Turbo 128K, and Doubao Seed 2.0 Lite) to detect text patterns matching the restricted content style.[4] Figure 1 shows that the monthly share of policy-targeted language remains persistently below 0.2% across all model classifications both before and after the campaign. This negligible baseline indicates that commercial websites had little room to adjust along the regulated margin. The evidence therefore points to the true locus of the policy response. The economically relevant regulatory shock fell mainly on financial content generated by independent self-media creators.

This localized impact makes China's TikTok (Douyin) a natural empirical setting for isolating the regulatory shock. As the country's leading short video platform, TikTok serves as

[3]China provides a useful institutional setting because public investment advice is explicitly regulated. Under the *Interim Measures for the Administration of Securities and Futures Investment Advisory Services*, only licensed professionals affiliated with qualified advisory institutions may provide public investment analysis, forecasts, and stock recommendations. This licensing rule provided a legal basis for the campaign. Before the campaign, many finfluencers blurred this boundary through recommendation-like commentary.

[4]We use the prompt-based screening procedure in Appendix Table A.1 to classify whether each article contains policy-targeted expressions corresponding to the eight categories identified in the campaign.

a central infrastructure for retail-facing financial influencers.[5] Unlike search-oriented channels, where visibility mainly reflects active information demand, TikTok amplifies market commentary through feed-based algorithmic curation. Platform-level enforcement data verify that the campaign imposed a binding constraint on this source of information supply. Within eight weeks of the intervention, TikTok issued eight public notices concerning financial accounts and penalized 1,684 financial accounts, including 1,357 independent creators that were permanently removed.

This targeted enforcement confirms that the campaign generated a direct contraction in finfluencer activity on China's largest short video platform. To translate this supply-side shock into causal evidence, our empirical strategy exploits cross-sectional variation in platform exposure. Crucially for our identification, TikTok's proprietary analytics provide granular stock-level visibility measures. We use these measures to capture each stock's pre-intervention exposure to finfluencer-related content. By comparing market quality outcomes for stocks with high versus low exposure, this framework allows us to estimate how restricting finfluencer activity affects intraday volatility and short-horizon price stability.

3 Data and Measurement

3.1 Sample Construction and Data Sources

Our data are drawn from multiple sources. The key explanatory variable, *Treat*, is constructed using the *TikTok Index* provided by Ocean Engine, a digital marketing and analytics platform operated by ByteDance, TikTok's parent company. It serves as a proxy for stock-level exposure to finfluencer activity prior to the CAC Finfluencer Campaign. Intraday volatility and other stock-level variables are obtained from the China Stock Market and Accounting Research (CSMAR) and Wind databases.

[5] Analysys Qianfan reports that China's TikTok had roughly 814.82 million quarterly active users in 2021Q2, ranking first among short video platforms. As of 2025, TikTok had more than 936 million monthly active users and over 200 million users engaging with finance-related content in China.

Our sample includes all A-share firms listed on China's main boards (Shanghai and Shenzhen). The sample period is centered on the August 27, 2021 regulatory shock and spans the six months before and after the announcement, as well as the month in which the policy was introduced. This yields a window from February 27, 2021 to March 26, 2022. For each stock-day observation within this period, we retain observations with non-missing key variables. After excluding observations with missing values in the primary dependent and control variables, the final sample contains 634,902 stock-day observations across 2,501 stocks. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles.

3.2 Measuring Intraday Volatility

Following Eaton et al. (2022), we measure intraday price volatility using five-minute intraday data. Our main dependent variables, Realized Volatility (RV) and Realized Range Volatility (RRV), are constructed from 5-minute price observations to capture short-horizon variation in stock prices around the CAC Finfluencer Campaign.

The first measure RV is defined as the sum of squared 5-minute log returns within each trading day. Specifically, for stock i on day t, RV is constructed as:

$$RV_{i,t} = \sum_{d=1}^{48} R_{i,t,d}^2, \quad R_{i,t,d} = 100 \times [\ln(I_{i,t,d}) - \ln(I_{i,t,d-1})] \tag{1}$$

where $I_{i,t,d}$ denotes the closing price of stock i within the 5-minute interval d on day t.

Our second measure is Realized Range Volatility (RRV), which is based on the intraday price range within each 5-minute interval. For each interval, we compute the log difference between the highest and lowest transaction prices and aggregate these values over the trading day. Specifically, for stock i on day t, RRV is defined as:

$$RRV_{i,t} = \sum_{d=1}^{48} \{100 \times [\ln(H_{i,t,d}) - \ln(L_{i,t,d})]\}^2 \tag{2}$$

where $H_{i,t,d}$ and $L_{i,t,d}$ denote the highest and lowest transaction prices of stock i within the 5-minute interval d on day t. Both $RV_{i,t}$ and $RRV_{i,t}$ are computed at the daily stock level and serve as proxies for short-term market instability and trading frictions.

3.3 Measuring Pre-Campaign Finfluencer Exposure

To facilitate causal identification, we distinguish stocks with high versus low finfluencer exposure on TikTok prior to the CAC Campaign. We use the *TikTok Index* as a proxy for this pre-campaign exposure. The *TikTok Index* is a keyword-level popularity measure that tracks the popularity of specific keywords over time and is commonly used to measure the intensity of content production and dissemination on TikTok. Using stock abbreviations as keywords, the resulting *TikTok Index* reflects the extent to which a stock is featured, discussed, and disseminated in TikTok content prior to the campaign.

A key platform feature that helps clarify what the *TikTok Index* captures is TikTok's feed-based, algorithmic content distribution. Unlike search-oriented platforms, where keyword popularity mainly reflects users' active information demand, TikTok exposes users to content primarily through personalized feeds rather than keyword search. As a result, the visibility of stock-related information on TikTok depends heavily on the continuous production and dissemination of such content and its amplification through user engagement.

In the financial domain, stock-related short videos may be produced by official accounts and media outlets. However, individual creators, especially finfluencers, represent an important source of this content and are particularly salient for retail investors. Consequently, on a content-centric platform such as TikTok, stock-level popularity is more likely to reflect the intensity of content supply and dissemination rather than investors' active information demand. This institutional feature therefore provides a natural basis for interpreting cross-sectional variation in the *TikTok Index* as differences in pre-campaign exposure to finfluencer-generated stock commentary.

According to the platform's documentation, the *TikTok Index* is constructed as a

weighted aggregation of multiple platform signals rather than a simple count of keyword mentions. It integrates information from content volume, dissemination intensity, and in-platform search activity associated with a given keyword and related content. On the content-centric platform TikTok, the index increases mechanically with the production and dissemination of stock-related content. In practice, it rises when creators mention or discuss a stock in short videos, when engagement through likes, comments, and shares amplifies the content, and when users search for firm-related keywords.[6]

Finfluencers produce stock-related short videos and contribute to the dissemination of such content through follower engagement and platform algorithms. Thus, variation in the *TikTok Index* plausibly reflects differences in stock-level exposure to finfluencer activity prior to the campaign. Higher index values therefore indicate greater pre-campaign exposure on TikTok. Following Xu et al. (2021), we next classify firms into treatment and control groups based on their pre-campaign exposure to finfluencer activity on TikTok.

For each stock, we compute the average *TikTok Index* over the trading days in the one-month period immediately preceding the CAC Finfluencer Campaign (July 27 to August 26, 2021).[7] Firms with an average index value greater than or equal to the sample median are assigned to the treatment group ($Treat = 1$), while those below the median form the control group ($Treat = 0$). This classification captures cross-sectional differences in stock visibility and attention on TikTok prior to the campaign.

3.4 Control Variables

Following prior studies on market quality and intraday price dynamics (Eaton et al., 2022; Dang et al., 2024), we include a set of firm-level control variables in our analysis, including *Size*, *Age*, *Ret*, and *ROE*. Specifically, *Size* is measured as the natural logarithm of a firm's

[6] The *TikTok Index* is similar to Google Search Volume in that both are platform-provided keyword popularity indices constructed using proprietary algorithms. Unlike Google Trends, however, the *TikTok Index* aggregates multiple platform signals rather than search activity alone. It also does not normalize keyword popularity by total platform search volume in the way Google Trends does.

[7] To ensure a stable measure of pre-campaign finfluencer exposure, we retain only stocks covered by the TikTok Index with non-missing values for at least 20 trading days during this period.

daily market capitalization of tradable shares, *Age* is defined as the natural logarithm of the number of years since the firm's establishment, *Ret* is the daily stock return calculated with cash dividends reinvested, and *ROE* is the firm's trailing twelve-month return on equity. Detailed variable definitions and notations are provided in Appendix Table A.2.

3.5 Summary Statistics

Table 1 reports summary statistics for the main variables, including the dependent variables (RV and RRV), the treatment indicator ($Treat$), the post-campaign indicator ($Post$), and the set of control variables. The summary statistics reveal substantial cross-sectional and time-series variation in intraday volatility across A-share firms. The mean value of RV is 7.80 with a standard deviation of 9.32, while the mean value of RRV is 19.38 with a standard deviation of 20.60. In terms of firm characteristics, the sample firms have an average market capitalization of RMB 21.81 billion and an average firm age of 24.08 years. The mean daily stock return is 0.07%, and the average return on equity (ROE) is 7.52%.

[INSERT TABLE 1 HERE]

4 Finfluencer Regulation and Market Quality

4.1 Baseline Difference-in-Differences Estimates

We begin our empirical analysis by examining how the 2021 CAC Campaign affects intraday stock price volatility. We estimate the following difference-in-differences specification:

$$IntraVol_{i,t} = \beta_0 + \beta_1 Treat_i \times Post_t + \boldsymbol{X}'\theta + \mu_i + \gamma_t + \varepsilon_{i,t} \tag{3}$$

where $IntraVol_{i,t}$ denotes the intraday volatility of stock i on day t. We measure intraday volatility using either Realized Volatility (RV) or Realized Range Volatility (RRV). The key independent variable is the interaction term $Treat_i \times Post_t$. The coefficient β_1 captures

the differential change in intraday volatility for stocks with high pre-campaign finfluencer exposure relative to stocks with low pre-campaign exposure after the CAC Finfluencer Campaign. μ_i and γ_t denote stock and date fixed effects, respectively. $\boldsymbol{X}$ includes the control variables described in Section 3.4, as well as the dependent variable measured on day $t-1$. $\varepsilon_{i,t}$ is the error term. If the CAC Finfluencer Campaign curtailed irrational trading driven by finfluencer promotion and improved the quality of information reaching retail investors, we expect $\hat{\beta}_1$ to be negative. Under this hypothesis, intraday price volatility should decline more for stocks with higher pre-campaign finfluencer exposure.

Table 2 reports the main results. Columns (1) and (2) use RV as the dependent variable, while columns (3) and (4) use RRV. All specifications include stock and date fixed effects, and columns (2) and (4) further control for firm-level characteristics. Standard errors, reported in parentheses, are clustered at the stock level in columns (1) and (3) and two-way clustered at both the stock and date levels in columns (2) and (4).

Across all specifications, the coefficient on $Treat \times Post$ is negative and statistically significant at the 1% level. This result indicates that stocks with greater pre-campaign exposure to finfluencer activity experience a larger decline in intraday volatility following the CAC Finfluencer Campaign.

The pattern holds for both RV and RRV. This consistency suggests that the campaign reduces short-term price fluctuations among treated stocks and thereby improves market quality. The results are also robust to the inclusion of controls for firm size, firm age, stock returns, and profitability. This robustness indicates that the effect is unlikely to be driven by changes in firm fundamentals or contemporaneous return dynamics.

[INSERT TABLE 2 HERE]

The estimated coefficients reported in columns (2) and (4) of Table 2 also indicate economically meaningful effects of the CAC Finfluencer Campaign on market quality. For Realized Volatility (RV), the coefficient on $Treat \times Post$ is -1.0433. The sample mean

of RV is 7.8027. This estimate implies that the campaign reduces intraday volatility by approximately 13.4% of its sample mean. For Realized Range Volatility (RRV), the corresponding coefficient is -2.0288, relative to a mean of 19.3826. This estimate represents a reduction in volatility magnitude of about 10.5% of its sample mean.

Thus, these results provide causal evidence of a reduction in intraday volatility following the CAC Finfluencer Campaign. By constraining finfluencer activity on TikTok, the campaign contributed to an improvement in overall market quality, particularly for stocks that were initially more exposed to finfluencer-driven attention.

4.2 Dynamic Effects

To examine the dynamic effects of the treatment and to assess the parallel trends assumption, we estimate an event-study version of our baseline DiD model.

Following Chen et al. (2023), we construct event-time indicators at the monthly frequency around the August 27, 2021 CAC Finfluencer Campaign. Relative time k ranges from -6 to $+6$. For ease of presentation, the periods from -6 to -4 are grouped into a single category labeled -4, and the periods from $+4$ to $+6$ are grouped into a single category labeled $+4$. The full sample period spans from February 27, 2021 to March 26, 2022. The month immediately following the campaign, from August 27 to September 26, 2021, is defined as the event month with $k = 0$. We define the indicator variable $D_{i,t}^{k}$ to equal one if stock i belongs to the treatment group and the observation falls within event time k, and zero otherwise. We then estimate the following regression to evaluate the dynamic treatment effects of the CAC Finfluencer Campaign:

$$IntraVol_{i,t} = \alpha + \sum_{k=-4,\ k\neq 0}^{4} \beta_k D_{i,t}^{k} + \boldsymbol{X}'\theta + \mu_i + \gamma_t + \varepsilon_{i,t} \tag{4}$$

where $IntraVol_{i,t} \in \{RV, RRV\}$ denotes stock i's intraday volatility on day t, $k = -4$ pools months -6 to -4, $k = 4$ pools months 4 to 6, and the event month $k = 0$ is omitted

as the reference period. The remaining empirical specifications follow the same setup as that in Equation (3).

[INSERT FIGURE 2 HERE]

Figure 2 plots the estimated coefficients for RV and RRV together with 95% confidence intervals. The coefficient in the event month is normalized to zero where $k = 0$. In the months leading up to the CAC Finfluencer Campaign where $k < 0$, the coefficients are small in magnitude and statistically indistinguishable from zero. This pattern indicates that treated and control firms exhibit similar pretrends in intraday volatility. After the campaign where $k > 0$, the coefficients become negative and statistically significant. This result shows a clear relative decline in both RV and RRV for high-exposure stocks. The reduction emerges quickly following the campaign and remains evident over several months. Overall, these patterns are consistent with a persistent improvement in short-horizon price stability for stocks with greater pre-campaign exposure to finfluencer activity.

4.3 Further Evidence on Causality

The event-study evidence reduces concerns about differential pre-trends, but two identification issues remain. The first is endogenous exposure. Stocks with high pre-campaign TikTok visibility may differ systematically from low-exposure stocks, and these differences may also shape post-campaign volatility dynamics. The second is measure interpretation. High TikTok exposure may capture general stock salience, conventional media attention, or demand for formal financial information rather than exposure to finfluencer-driven content. We address these issues in turn.

We begin with the first concern: pre-campaign TikTok exposure may be endogenously related to stock characteristics that also affect post-campaign volatility dynamics. Table 3 reports four sets of tests.

First, we use propensity score matching to construct a treatment-control sample with stronger overlap in observable characteristics. The propensity score model includes the baseline firm controls, industry fixed effects based on the 2012 CSRC classification, and province fixed effects. We match treated stocks to control stocks using nearest-neighbor matching within a caliper of 0.01 and restrict the sample to common support. Figure 3 shows that matching substantially improves the overlap in estimated propensity scores. Columns (1) and (2) of Panel A show that, in the matched sample, the coefficient on $Treat \times Post$ remains negative and statistically significant for both RV and RRV.

[INSERT FIGURE 3 HERE]

Second, to further address observable covariate imbalance without discarding observations, we use entropy balancing as an alternative reweighting approach. This method reweights control observations so that the first three moments of the covariates match those of the treated group while retaining the full sample. Columns (3) and (4) of Panel A show that the estimates remain negative and statistically significant after applying these weights.

Third, to address concerns that the baseline estimates depend on a restrictive linear specification, we use double debiased machine learning to allow for more flexible confounding structures. We first residualize the outcome, the treatment variable, and the controls with respect to stock and date fixed effects. We then estimate the nuisance functions in a partial linear model using 10-fold cross-fitting with Random Forest and Lasso learners. Panel B shows that the estimated treatment effect remains negative and statistically significant across both learners and both volatility measures. These results reduce concerns that the baseline estimates are driven by a particular linear specification.

Finally, to assess sensitivity to omitted variable bias, we use the Impact Threshold for a Confounding Variable (ITCV) and Robustness of Inference to Replacement (RIR) methods. Panel C shows that an omitted variable would need a joint correlation of about 0.013 for RV and 0.014 for RRV with both treatment and outcome to eliminate statistical

significance. Except for the lagged dependent variables, these thresholds exceed the corresponding impacts of the observed covariates in the baseline model. Panel D shows that more than 84% of the sample would need to be replaced by zero effect observations for the estimates to lose significance.

[INSERT TABLE 3 HERE]

These tests reduce the concern that the baseline effect is driven by endogenous exposure related to observable imbalance, functional form, or modest omitted-variable bias. We next examine whether TikTok exposure captures creator-driven information supply rather than general stock salience. A remaining concern is that high pre-campaign TikTok exposure may simply identify stocks that were already popular, heavily covered, or temporarily salient before the campaign. Under this alternative interpretation, the post-campaign decline in volatility could reflect mean reversion among salient stocks rather than reduced exposure to finfluencer-related content.

Table 4 examines this alternative explanation. Panel A first addresses the possibility that treatment status reflects a short-lived spike in TikTok popularity preceding the campaign. Such transitory attention could mechanically predict later mean reversion in volatility. We therefore redefine treatment using average TikTok exposure over the 3 and 6 months before the campaign, rather than the baseline 1-month window. The coefficient on $Treat \times Post$ remains negative and statistically significant for both RV and RRV. The baseline result is therefore unlikely to be driven by transitory pre-campaign attention.

Panel B addresses the concern that TikTok exposure simply captures general stock salience. If this explanation were correct, stocks with high pre-campaign conventional financial news coverage should exhibit a similar post-campaign decline in volatility. They do not. The coefficient on $PseudoTreat \times Post$ is statistically insignificant. When the newspaper-based placebo and the TikTok-based treatment enter the regression jointly, the coefficient on $Treat \times Post$ remains negative and significant.

Panel C further separates TikTok exposure from demand for financial information. This test addresses the possibility that TikTok exposure proxies for investors' general information demand rather than exposure to finfluencer-driven content. We residualize TikTok exposure using Eastmoney-based measures of news reading, announcement reading, and analyst report reading. The residualized treatment remains negatively and significantly associated with post-campaign volatility. This evidence indicates that the TikTok-based treatment contains variation beyond observable demand for formal financial information.

[INSERT TABLE 4 HERE]

These results strengthen the interpretation of the TikTok-based exposure measure. The baseline effect is not driven by transitory pre-campaign popularity, nor is it replicated by conventional financial news coverage. It also remains after removing observable demand for formal financial information from the exposure measure. The evidence therefore supports the view that TikTok exposure captures variation in creator-driven financial information supply rather than general stock salience.

5 Mechanisms

5.1 Information Quality and Belief Alignment

Several studies link the information environment to market quality, particularly through high-frequency price dynamics. Intraday volatility reflects the arrival of information and the trading-related adjustment of prices during market hours. Cleaner and more reliable information flows should therefore reduce short-horizon noise in returns (French and Roll, 1986). More broadly, improvements in disclosure and reductions in information asymmetry are associated with better market functioning. In contrast, heterogeneous interpretations of common signals can generate disagreement, trading, and amplified price fluctuations. These findings suggest that attenuating noisy or biased signals can stabilize prices (Diamond and

Verrecchia, 1991; Harris and Raviv, 1993). Related work on information dissemination further shows that shifts in information supply and visibility can causally affect investor trading behavior (Engelberg and Parsons, 2011).

The CAC Finfluencer Campaign provides a setting in which enforcement targeted the production and dissemination of misleading or biased financial content by unlicensed individual creators. This feature motivates an information-content mechanism. If the crackdown reduced the prevalence and salience of low-quality stock promotion, the retail information environment should become more informative and more coherent. In the remainder of this subsection, we evaluate this mechanism in two steps. We first examine whether the campaign reshaped finfluencers' content composition. We then study whether these supply-side changes are accompanied by improvements in the quality of stock-related signals and greater alignment in retail beliefs.

We begin by examining how finfluencers adjusted their posting behavior around the CAC Finfluencer Campaign. Using Ocean Engine's similar-creator recommendation function, we expand a seed list of finfluencer accounts through three iterative rounds. This procedure allows us to construct a broad sample of finfluencers that aligns with the platform's own similarity and discovery structure.[8] Based on this account set, we collect finfluencer posts published on TikTok around the campaign, yielding 66,394 posts from 1,972 finfluencer accounts.

To characterize the informational content of these posts, we implement a multi-label classification procedure using the reasoning large language model (DeepSeek-R1), following related applications in the disclosure and information environment literature (Bertomeu et al., 2025). Each post is classified into one or more content dimensions: *Stock*, *Industry*,

[8]We first identify a seed list of finfluencer accounts using keyword-based creator search on Ocean Engine, including the terms "stock trading," "finance," and "financial markets." For each keyword, the platform returns the top 30 most relevant creators. For each seed finfluencer, Ocean Engine further recommends up to 20 similar creators based on content and audience similarity. We iteratively apply this similar-creator expansion procedure for three rounds. After removing duplicates and restricting the sample to finfluencers who posted at least once during the period from July 27 to September 26, 2021, we obtain a final sample of 1,972 finfluencer accounts. The same search and expansion procedure can be replicated at https://trendinsight.oceanengine.com/arithmetic-index?type=3.

Market, and *Macro*, depending on whether the post contains information related to the corresponding category. The classification prompt is reported in Appendix Table A.3. We then examine how content composition changes around the intervention by comparing pre- versus post-campaign posting behavior within finfluencers while controlling for finfluencer fixed effects.

Panel A of Table 5 shows a clear post-campaign shift in finfluencer content. Stock-specific posts become less frequent, while industry-level discussions become more common. We find no meaningful change in market-level or macro-level content. In economic terms, relative to the sample mean, the likelihood that a post contains stock-specific content declines by 11.43%, while the likelihood that it contains industry-level content increases by 1.04%. This pattern is consistent with enforcement discouraging unqualified stock promotion and recommendation-like content. It also provides platform-level evidence that the campaign reshaped the supply of finfluencer-generated financial information along the targeted margin. These results set the stage for our subsequent tests on the retail information environment and market quality.

We next examine whether the observed shift in finfluencer content supply translates into improvements in the retail-facing information environment. We focus on two complementary outcomes that capture the revealed quality of stock-related signals and the degree of alignment in retail beliefs. Specifically, we estimate the following stock-day difference-in-differences specification:

$$Y_{i,t} = \beta_0 + \beta_1 Treat_i \times Post_t + \boldsymbol{X}'\theta + \mu_i + \gamma_t + \varepsilon_{i,t} \tag{5}$$

where $Y_{i,t}$ denotes either $PostQuality_{i,t}$ or the Sentiment Conformity Index $SCI_{i,t}$. We measure $PostQuality_{i,t}$ as the average number of likes per stock-related post on the East-money Guba forum for stock i on day t. This measure captures engagement-weighted signal quality. It reflects which messages retail investors perceive as sufficiently useful or

credible to endorse and propagate, rather than the sheer volume of information produced. We obtain the Sentiment Conformity Index $SCI_{i,t}$ from CSMAR. All specifications include stock and date fixed effects, as well as standard firm-level controls.

Panel B of Table 5 reports the results. When $PostQuality_{i,t}$ is the dependent variable, the coefficient on $Treat_i \times Post_t$ is positive. This finding indicates that after the campaign, high-exposure stocks experience a larger increase in engagement-weighted post quality. This pattern is consistent with the intervention reducing the prevalence of low-quality stock promotion and shifting investor engagement toward more informative stock-related discussion. It is unlikely to reflect a mechanical change in posting volume. When $SCI_{i,t}$ is the dependent variable, the coefficient on $Treat_i \times Post_t$ is also positive. This result indicates higher sentiment conformity and lower retail disagreement for high-exposure stocks after the campaign. A decline in disagreement is consistent with a less fragmented retail information set and with fewer noisy or biased signals generating divergent interpretations.

[INSERT TABLE 5 HERE]

Overall, the evidence in this subsection is consistent with an information-environment channel. The campaign appears to discourage stock-specific promotion and shift finfluencer content toward broader, less recommendation-like discussion. High-exposure stocks also experience improvements in engagement-weighted post quality and stronger sentiment conformity among retail investors. These patterns suggest that the CAC Finfluencer Campaign improved the retail information environment by reducing noisy stock-level promotion and making retail beliefs more coherent.

5.2 Retail Attention and Participation

Beyond the information-content channel examined in Section 5.1, the CAC Finfluencer Campaign may also affect market quality through an attention and participation channel. Although this channel is not the policy's primary objective, it may naturally arise when

highly engaging content is curtailed on a feed-based platform. Reduced finfluencer activity can lower stock-level visibility on TikTok and dampen attention-driven retail participation. A strand of literature links attention-driven retail trading to excess short-horizon volatility and weaker market quality (Barber and Odean, 2000; Foucault et al., 2011; Eaton et al., 2022). In this section, we evaluate this complementary mechanism by testing whether high-exposure stocks experience larger post-campaign declines in platform visibility, as well as corresponding declines in off-platform retail discussion and retail trading intensity.

We first examine how the regulatory intervention affected stock-level popularity on TikTok. Specifically, we estimate the following stock-day difference-in-differences specification within a one-month window surrounding the campaign (July 27 to September 26, 2021):

$$Index_{i,t} = \beta_0 + \beta_1 Treat_i \times Post_t + \boldsymbol{X}'\theta + \mu_i + \gamma_t + \varepsilon_{i,t} \tag{6}$$

where $Index_{i,t}$ denotes either $\log(1 + \textit{TikTok Index})$ or $\log(1 + \textit{TikTok Search})$ for stock i on day t (Lee and So, 2017). $Treat_i$ equals one for stocks with pre-campaign exposure to finfluencer activity on TikTok at or above the sample median. $Post_t$ equals one for dates on or after August 27, 2021. μ_i and γ_t denote stock and date fixed effects, respectively. $\boldsymbol{X}$ includes firm size and lagged stock returns. We also control for the lagged log-transformed rolling 20-trading-day average of the corresponding TikTok measure.

Panel A of Table 6 reports the results. Across specifications, the coefficient on $Treat_i \times Post_t$ is negative. This finding indicates that high-exposure stocks experience larger post-campaign declines in TikTok visibility. The estimates in Columns (2) and (4) imply declines of about 7.54% and 9.09%, respectively, relative to low-exposure stocks. These patterns provide direct evidence that the campaign materially reduced the platform visibility of stocks that had been more strongly exposed to finfluencer activity prior to the intervention.

Beyond platform visibility, retail investors also reveal engagement with stock information through online discussion forums. Activity on investor communities provides a useful

proxy for retail participation in stock-specific discourse and related sentiment. Prior work shows that such activity is closely connected to market outcomes such as trading activity and volatility (Antweiler and Frank, 2004; Chen et al., 2014; Cookson and Niessner, 2020). If the campaign dampened highly salient finfluencer content, retail investors may have become less inclined to comment on and debate the affected stocks in these forums. To evaluate this margin, we extend the baseline difference-in-differences framework in Equation (5) by replacing the outcome variable with stock-day measures of retail discussion and trading activity. All specifications include stock and date fixed effects, as well as the same set of firm-level controls.

We first measure retail discussion using $GubaPost_{i,t}$, defined as the logarithm of one plus the number of firm-specific posts on the Eastmoney Guba forum for stock i on day t. Eastmoney Guba is a major retail investor forum in China and is analogous to large online investor discussion boards in U.S. markets. Panel B of Table 6 reports the results. The coefficient on $Treat_i \times Post_t$ is negative, indicating that high-exposure stocks experience larger declines in forum posting after the campaign. The estimate in Column (2) implies a reduction of about 3.76% relative to low-exposure stocks. This pattern is consistent with diminished retail engagement with these stocks in off-platform discussion. Prior evidence links message volume to trading intensity and volatility (Antweiler and Frank, 2004; Chen et al., 2014). The decline in retail opinion expression therefore represents a meaningful reduction in retail attention.

[INSERT TABLE 6 HERE]

We next examine whether the decline in retail engagement is mirrored in retail trading activity. Here, $Y_{i,t}$ is $RetailTrade_{i,t}$, defined as the logarithm of daily trading volume generated by small orders below RMB 40,000 using transaction-level data from Wind. This order-size–based measure is commonly used to proxy for the intensity of retail investor participation (Barber and Odean, 2000; Eaton et al., 2022). Columns (3) and (4) of Panel B of

Table 6 report the results. The coefficient on $Treat_i \times Post_t$ is negative and statistically significant. This finding indicates that high-exposure stocks experience larger post-campaign declines in retail trading intensity. The estimate in Column (4) implies a reduction of approximately 4.77% relative to low-exposure stocks.

Overall, these results suggest that the campaign dampened attention-driven retail participation for stocks that were more exposed to finfluencers prior to the intervention. This provides a complementary channel consistent with the observed decline in intraday volatility and the improvement in market quality.

6 Cross-Sectional Heterogeneity and Robustness

6.1 Heterogeneity Analysis

6.1.1 The Role of Information Environment

Previous literature emphasizes that a firm's information environment, defined by the amount, quality, and timeliness of value-relevant information available to investors, plays a central role in shaping asset prices and capital market outcomes.

Firms with more transparent and informative environments exhibit higher valuations, lower information asymmetry, and lower costs of capital. Their stock prices also tend to reflect firm fundamentals more closely rather than noise or opacity (Leuz and Verrecchia, 2000; Lang et al., 2003; Durnev et al., 2003; Jin and Myers, 2006). In our setting, this literature implies that the effect of the CAC Finfluencer Campaign on intraday volatility should vary with the strength of a firm's existing information environment, namely how strongly investors can rely on established and regular information channels.

To examine this prediction, we capture cross-sectional variation in firms' information environments using three standard proxies: analyst coverage, institutional ownership, and disclosure quality.

[**INSERT TABLE 7 HERE**]

We first examine heterogeneity based on analyst coverage as a proxy for the strength of the information environment. Analyst coverage is measured as the number of analyst reports issued for a firm over the past year. It is commonly viewed as capturing the extent to which professional analysts supply monitoring, interpretation, and firm-specific information to the market (Andrade et al., 2013; Lee and So, 2017). We define *Analyst* as an indicator equal to one when a firm's analyst coverage is at or above the cross-sectional median and zero otherwise, thereby classifying firms into information-rich and information-poor groups. Panel A of Table 7 reports the results. In both the *RV* and *RRV* specifications, the interaction term $Treat \times Post \times Analyst$ is positive and statistically significant at the 5% level. This finding indicates that the decline in intraday volatility following the CAC Finfluencer Campaign is more pronounced for firms with lower analyst coverage. When analyst monitoring is limited, retail investors rely more heavily on informal information channels and are therefore more exposed to finfluencer commentary. As a result, restricting finfluencer activity yields a larger improvement in short-horizon market quality for these information-poor firms.

We next examine heterogeneity based on institutional ownership as a second proxy for the information environment. Institutional ownership is measured as the fraction of tradable A-shares held by institutional investors at the most recent quarter end. It is commonly interpreted as reflecting the presence of informed and sophisticated market participants who supply research and improve price discovery (Boehmer and Kelley, 2009). We define *Institution* as an indicator equal to one when institutional ownership is at or above the cross-sectional median and zero otherwise. Panel B of Table 7 documents similar results when firms are grouped by institutional ownership. The interaction term $Treat \times Post \times Institution$ is positive and statistically significant at the 5% level. This result indicates that the decline in intraday volatility following the CAC Finfluencer Campaign is more pronounced for firms with lower institutional ownership. Institutional investors

are typically better informed and more sophisticated, which can stabilize the trading environment and reduce the influence of retail-driven noise. Accordingly, when institutional participation is limited, finfluencer-driven attention and sentiment are likely to exert greater influence on prices prior to the regulation, leading to a sharper volatility decline afterward.

Finally, we examine heterogeneity based on disclosure quality as a third proxy for the information environment. Disclosure quality is proxied by the corporate transparency ratings issued by the Shanghai and Shenzhen Stock Exchanges, which evaluate firms' reporting transparency and information disclosure practices. Prior research treats disclosure quality as a central dimension of the information environment and shows that more transparent firms face lower information asymmetry and exhibit different patterns of informed trading and crash risk (Jin and Myers, 2006; Maffett, 2012). We define $Transparency$ as an indicator equal to one if a firm receives an A rating in the most recent transparency evaluation and zero otherwise. Panel C of Table 7 reports heterogeneity by disclosure quality using the exchanges' transparency ratings. The interaction term $Treat \times Post \times Transparency$ is positive and statistically significant at the 1% level. This finding indicates that the decline in intraday volatility following the CAC Finfluencer Campaign is more pronounced for firms with lower disclosure quality. When firms provide less transparent information, investors face greater informational frictions and are more likely to rely on finfluencer commentary as a supplementary source of signals. As a result, restricting such content yields larger improvements in short-horizon market quality for these information-poor firms.

The evidence across all three dimensions points to a unified conclusion: the effect of finfluencer regulation on market quality is stronger in information-poor environments. The CAC Finfluencer Campaign generates the largest reductions in intraday volatility among firms with limited analyst coverage, low institutional ownership, and weak disclosure quality, where investors face fewer reliable information sources. These findings reinforce the interpretation that finfluencers act as substitutes for formal information channels. Restricting such content therefore disproportionately improves market quality in settings characterized

by larger information gaps.

6.1.2 The Role of Audience Gender

Prior research documents systematic and persistent differences in investment behavior across gender groups. Male investors trade more frequently, hold less diversified portfolios, and exhibit greater overconfidence. These behaviors often reduce net returns (Barber and Odean, 2001; Anginer et al., 2025; Edelen et al., 2025). Psychological and behavioral evidence further suggests that men, especially younger men, respond more strongly to attention shocks and take greater risks than women. This pattern is consistent with evidence that men are more susceptible to external stimuli (Arnold et al., 2022).

[INSERT TABLE 8 HERE]

While previous literature emphasizes the importance of investor heterogeneity in shaping attention and trading responses, existing attention-based measures rarely provide information on the demographic composition of the underlying audience. In our setting, the *TikTok Index* reports not only aggregate keyword popularity but also audience-profile statistics, including the gender distribution of users interacting with each keyword. These gender-specific exposure measures allow us to track, on a daily basis, the share of male users associated with each stock-specific keyword. This feature enables a direct test of whether the regulatory shock differentially affects stocks whose finfluencer audiences are more male-dominated.

Table 8 incorporates this demographic information by interacting the treatment with the share of male users, denoted $MaleRatio$. Across all specifications, the coefficient on $Treat \times Post \times MaleRatio$ is negative and statistically significant at the 1% level. This result indicates that the improvement in market quality following the regulation is stronger for stocks whose finfluencer-related audience is more male-dominated. This pattern is consistent with prior evidence that men trade more actively, exhibit greater overconfidence,

and respond more strongly to attention shocks (Barber and Odean, 2001; Arnold et al., 2022). Restricting finfluencer activity therefore dampens speculative trading more strongly in stocks with male-dominated audiences, leading to a larger reduction in intraday volatility. Given that male investors account for the majority of retail trades in China (Jones et al., 2025), these results highlight gender composition as an important channel through which finfluencer regulation affects short-horizon market quality.

6.2 Robustness Tests

We perform a series of additional robustness analyses to assess the sensitivity of our results to alternative sample constructions, variable definitions, and empirical specifications.

Robustness to Search-Demand-Based Explanations A remaining concern is that our main results may be driven by active search demand on TikTok rather than by finfluencer-related content production and dissemination. Since the *TikTok Index* aggregates multiple components, including content production, dissemination intensity, and in-platform search activity, it is possible that variation in the index partly reflects investors' active information demand rather than exposure to finfluencer-generated content. To address this concern and adopt a more conservative identification strategy, we isolate the component of platform popularity that is orthogonal to search-driven attention.[9]

Specifically, we residualize the *TikTok Index* on *TikTok Search* using rolling regressions and retain only the residual component. This residual captures variation in platform popularity that is not explained by contemporaneous search activity. The residual-based measure therefore reflects the portion of TikTok popularity driven primarily by content production and dissemination rather than by user search demand. Based on this orthogonalized measure, we reclassify stocks into treatment and control groups using pre-campaign

[9]Consistent with this interpretation, Appendix Table A.4 shows that the *TikTok Index* and *TikTok Search* exhibit distinct intertemporal dynamics. While contemporaneous values of both measures are correlated, changes in the *TikTok Index* strongly predict future search activity, whereas the reverse relationship is considerably weaker. This pattern suggests that content-driven platform popularity tends to lead, rather than merely reflect, search-based attention.

residualized TikTok popularity and re-estimate our baseline DiD specifications.

Panel A of Table 9 reports the results. Columns (1) and (2) use Realized Volatility (RV) as the dependent variable, while Columns (3) and (4) use Realized Range Volatility (RRV). Across all specifications, the coefficient on the residual-based treatment indicator interacted with the post-campaign dummy remains negative and statistically significant. These results indicate that even after removing the component of platform popularity attributable to active search behavior, stocks with higher exposure to finfluencer-related content continue to experience a larger decline in intraday volatility following the CAC Finfluencer Campaign. This evidence alleviates concerns that our main findings are driven purely by changes in active search behavior on the platform. Instead, the results support the interpretation that the regulation-induced reduction in finfluencer content production and dissemination plays a central role in shaping the observed market response.

Excluding Consumer and Brand-Related Firms. A potential concern with attention-based measures constructed from keyword popularity is that part of the measured activity may reflect non-investment-related interest rather than stock-specific attention. Prior studies show that search-based attention measures may include irrelevant searches because some keywords are also used outside financial contexts (deHaan et al., 2025).[10] In our setting, this concern mainly arises because some firm names correspond to well-known consumer companies or widely recognized brands. For instance, searches for Guizhou Moutai may capture interest in the liquor product rather than attention related to stock trading. To mitigate this concern, we re-estimate our baseline specifications after excluding firms in consumer- and brand-related industries as defined by the CSRC 2012 industry classification.[11] Panel B of Table 9 reports the results. Across both volatility measures, the

[10] deHaan et al. (2025) discuss that search indices constructed from stock tickers can suffer from keyword ambiguity, as the same term may refer to non-financial objects (e.g., "CAT" may reflect searches for animals rather than Caterpillar Inc.), causing unrelated queries to be included in the index.

[11] The excluded categories include food manufacturing, alcohol beverage and refined tea manufacturing, catering, retail, accommodation, cultural educational arts sports and recreational products manufacturing, textile apparel and accessories manufacturing, leather fur feather products and footwear manufacturing, furniture manufacturing, culture and arts, news and publishing, radio television film and audiovisual production, and automobile manufacturing.

coefficient on $Treat \times Post$ remains negative and statistically significant at the 1% level. This finding indicates that the documented post-campaign decline in intraday volatility is not driven by measurement noise related to consumer-brand discussions.

Additional Evidence Using Toutiao Popularity. Because the CAC Finfluencer Campaign targeted finfluencer activity broadly across major online content platforms, an important question is whether our results extend beyond TikTok-based exposure. To address this issue, we re-estimate our baseline specifications using an alternative platform-based measure from Toutiao. Toutiao is a major Chinese content platform that aggregates news, including financial and market-related content, and reaches a large user base. Specifically, we replace the TikTok-based treatment classification with a Toutiao-based measure by defining *Toutiao Popular Stocks* using the pre-campaign *Toutiao Index*. The indicator $ToutiaoHot$ equals one if a stock's average *Toutiao Index* over the pre-campaign window from July 27 to August 26, 2021 is at or above the cross-sectional median and zero otherwise. As reported in Panel C of Table 9, the coefficient on $ToutiaoHot \times Post$ is negative and statistically significant at the 1% level across both intraday volatility measures. This result indicates that finfluencer regulation is associated with improved market quality even when exposure is measured using activity on other major online content platforms.

[INSERT TABLE 9 HERE]

Additional Evidence from Bid-Ask Spreads. In addition to intraday volatility, we examine whether the CAC Finfluencer Campaign improved market quality by assessing its effect on bid-ask spreads. Following Eaton et al. (2022), we use two transaction-level measures of bid-ask spreads obtained from the CSMAR database. The first measure, $Spread^{Volume}$, is the volume-weighted average of the effective bid-ask spread. The second measure, $Spread^{Amount}$, is the amount-weighted average of the effective bid-ask spread. We then re-estimate the baseline DiD specification using these spread measures as the dependent variables. Table 10 reports the results. Columns (1) and (2) use $Spread^{Volume}$,

while columns (3) and (4) use $Spread^{Amount}$. Across both measures and specifications, the coefficient on $Treat \times Post$ is negative and statistically significant at the 5% level. This finding indicates that treated stocks experienced a post-campaign narrowing of effective spreads, consistent with an improvement in market quality following the campaign.[12]

[INSERT TABLE 10 HERE]

Other Robustness Tests. Table 11 reports additional sensitivity tests. Panel A varies the empirical specification. We replace the binary treatment indicator with the log-transformed mean of the pre-campaign *TikTok Index*, and we add industry-month fixed effects based on the 2012 CSRC classification to absorb time-varying industry shocks. Panel B varies the treatment construction. We define pre-campaign exposure using the median rather than the mean of the *TikTok Index*, and we compare stocks in the top and bottom terciles of pre-campaign exposure. Across these specifications, the estimated post-campaign effect remains negative and statistically significant. The results therefore do not appear to depend on the binary treatment definition, the mean-based exposure measure, the median-split cutoff, or the baseline fixed-effect structure.[13]

[INSERT TABLE 11 HERE]

Placebo Test. To rule out the possibility that our baseline estimates are driven by unobserved factors unrelated to the CAC Finfluencer Campaign, we conduct a placebo test based on randomized treatment assignment. We randomly reassign stocks to placebo treatment and control groups while preserving the original group sizes. In each simulation, we randomly assign 1,253 stocks to the placebo treatment group and 1,248 stocks to the placebo control group, for a total of 2,501 stocks. We then re-estimate the baseline

[12] To complement the baseline difference-in-differences estimates, we further examine dynamic effects for bid-ask spreads in Appendix Figure A.1. The estimated coefficients show no systematic pre-trend between treated and control firms and a clear post-campaign decline in spreads, supporting the parallel trends assumption for this alternative measure of market quality.

[13] Additional robustness checks are reported in Appendix Table A.5. These tests include non-trading days in the exposure measure, use top-versus-bottom decile groupings, and extend the post-event horizon to nine and twelve months. The results remain similar.

difference-in-differences specification to obtain the corresponding placebo coefficient on the treatment interaction. We repeat this procedure 500 times to generate distributions of placebo estimates for RV and RRV. Figure 4 plots the resulting coefficient distributions under random assignment. The placebo estimates are centered close to zero and display approximately symmetric distributions, indicating that randomly assigned treatment labels do not generate systematic volatility effects. In contrast, the actual treatment estimates reported in Table 2 of -1.0433 for RV and -2.0288 for RRV lie far in the left tail of the placebo distributions. This pattern suggests that the observed post-campaign decline in intraday volatility reflects a causal effect of the CAC Finfluencer Campaign rather than random variation or unobserved confounding factors.

[INSERT FIGURE 4 HERE]

7 Conclusion

Finfluencers operate as highly visible intermediaries within the modern retail information environment. Yet little is known about how their platform-mediated activity affects aggregate stock market quality. Creator-driven commentary may improve information processing, but it may also amplify speculative noise. We examine this tension using the 2021 Clean and Orderly Cyberspace campaign in China. This regulatory intervention tightened platform enforcement and sharply curtailed the supply of unregulated financial advice. We isolate this shock using China's TikTok, which provides proprietary stock-level visibility measures that allow us to capture exposure to creator-driven financial content.

Using A-share intraday data, we classify stocks by their pre-intervention TikTok exposure. After the regulatory shock, stocks with higher baseline finfluencer exposure experience a pronounced and sustained decline in short-horizon price instability. The reduction in realized volatility is economically meaningful. The baseline findings remain robust after addressing endogenous exposure, flexible confounding structures, and omitted-variable

sensitivity. The effect also persists after separating platform-specific visibility from conventional news coverage and active demand for formal financial information. The evidence indicates that the volatility decline reflects a contraction in creator-driven information supply rather than a general decline in stock salience.

We trace this market improvement to two complementary mechanisms. The first operates through the retail information environment. After the intervention, content supply shifts away from stock-specific commentary toward broader industry narratives. High-exposure stocks also exhibit higher engagement-weighted post quality and stronger sentiment conformity, consistent with a more coherent retail information environment. The second mechanism operates through investor attention and trading. The contraction in platform visibility reduces retail engagement and dampens small-order trading intensity. These channels suggest that the policy stabilized prices by reducing noisy stock-level promotion and limiting the entry of attention-driven retail traders.

Cross-sectional heterogeneity clarifies where this effect is strongest. The volatility reduction is concentrated among firms with weaker formal information environments. When professional information intermediaries are limited, retail investors rely more heavily on informal social signals, leaving prices more exposed to platform-amplified noise. Audience demographics provide further support for this interpretation. The stabilizing effect is stronger for stocks with more male-dominated finfluencer audiences, where platform content is more likely to translate into aggressive retail trading.

These findings identify an upstream source of market quality in modern retail finance. Digital platforms do not merely reflect investor attention. They shape which financial narratives reach retail audiences and how strongly those narratives are amplified. When retail participation depends heavily on informal social signals, regulation of algorithmic information supply can affect price stability. Platform governance has therefore become an important component of market governance. Understanding how feed-based content systems curate retail attention is essential for studying price formation in the digital era.

Acknowledgments

Qingfu Liu acknowledges the support of the National Natural Science Foundation of China (72595845, 72121002, and 72595840). Qingfu Liu thanks the Shanghai Institute for Mathematics and Interdisciplinary Science (SIMIS) for their financial support. This research was funded by SIMIS under grant number [SIMIS-ID-2025-TF]. The other authors received no financial support for this research. The authors have no financial relationships or other potential conflicts of interest to disclose.

References

Andrade, S. C., J. Bian, and T. R. Burch (2013). Analyst coverage, information, and bubbles. *Journal of Financial and Quantitative Analysis 48*(5), 1573–1605.

Anginer, D., S. X. Han, and Ç. Yıldızhan (2025). Do individual investors ignore transaction costs? *Journal of Financial and Quantitative Analysis 60*(8), 3899–3931.

Antweiler, W. and M. Z. Frank (2004). Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance 59*(3), 1259–1294.

Arnold, M., M. Pelster, and M. G. Subrahmanyam (2022). Attention triggers and investors' risk-taking. *Journal of Financial Economics 143*(2), 846–875.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance 55*(2), 773–806.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

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

Bertomeu, J., Y. Lin, Y. Liu, and Z. Ni (2025). The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. *Journal of Accounting and Economics 80*(1), 101782.

Bhargava, H. K. (2022). The creator economy: Managing ecosystem supply, revenue sharing, and platform design. *Management Science 68*(7), 5233–5251.

Blankespoor, E., G. S. Miller, and H. D. White (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *The Accounting Review 89*(1), 79–112.

Boehmer, E. and E. K. Kelley (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies 22*(9), 3563–3594.

Chen, H., Z. Chen, Z. He, J. Liu, and R. Xie (2023). Pledgeability and asset prices: Evidence from the Chinese corporate bond markets. *The Journal of Finance 78*(5), 2563–2620.

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

Comerton-Forde, C., V. Grégoire, and Z. Zhong (2019). Inverted fee structures, tick size, and market quality. *Journal of Financial Economics 134*(1), 141–164.

Cookson, J. A., J. E. Engelberg, and W. Mullins (2023). Echo chambers. *The Review of Financial Studies 36*(2), 450–500.

Cookson, J. A., C. Fox, J. Gil-Bazo, J.-P. Imbet, and C. Schiller (2026). Social media as a bank run catalyst. *Journal of Financial Economics 176*, 104218.

Cookson, J. A., W. Mullins, and M. Niessner (2024). Social media and finance. Available at SSRN.

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

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

Dang, T. V., W. Li, and Y. Wang (2024). Government stock purchase undermines price informativeness: Evidence from China's national team. *Journal of Financial and Quantitative Analysis 59*(5), 2340–2374.

deHaan, E., A. Lawrence, and R. Litjens (2025). Measuring investor attention using Google search. *Management Science 71*(7), 6275–6297.

Diamond, D. W. and R. E. Verrecchia (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance 46*(4), 1325–1359.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin (2003). Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research 41*(5), 797–836.

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

Edelen, R. M., K. Y. L. Fong, and J. Han (2025). Regulating inattention in fee-based financial advice. *Journal of Financial Economics 164*, 103985.

Engelberg, J. E. and C. A. Parsons (2011). The causal impact of media in financial markets. *The Journal of Finance 66*(1), 67–97.

Foucault, T., D. Sraer, and D. J. Thesmar (2011). Individual investors and volatility. *The Journal of Finance 66*(4), 1369–1406.

French, K. R. and R. Roll (1986). Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics 17*(1), 5–26.

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

Hull, I. and Y. Qi (2024). The impact of finfluencers on retail investment. Available at SSRN.

Jin, L. and S. C. Myers (2006). R^2 around the world: New theory and new tests. *Journal of Financial Economics 79*(2), 257–292.

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

Kakhbod, A., S. M. Kazempour, D. Livdan, and N. Schuerhoff (2023, July). Finfluencers. Research Paper 23-30, Swiss Finance Institute.

Kogan, S., T. J. Moskowitz, and M. Niessner (2023). Social media and financial news manipulation. *Review of Finance 27*(4), 1229–1268.

Lang, M. H., K. V. Lins, and D. P. Miller (2003). ADRs, analysts, and accuracy: Does cross-listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research 41*(2), 317–345.

Lee, C. M. C. and E. C. So (2017). Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics 124*(2), 331–348.

Leuz, C. and R. E. Verrecchia (2000). The economic consequences of increased disclosure. *Journal of Accounting Research 38*(Supplement), 91–124.

Liang, Y., X. Chen, S. Han, J. Zhang, and Y. Chen (2025). Is the money spent on short-form video social platforms worth it? the role of advertising spillover in a large-scale randomized field experiment on ByteDance. *Marketing Science 44*(5), 1125–1144.

Maffett, M. (2012). Financial reporting opacity and informed trading by international institutional investors. *Journal of Accounting and Economics 54*(2–3), 201–220.

Malinova, K. and A. Park (2015). Subsidizing liquidity: The impact of make/take fees on market quality. *The Journal of Finance 70*(2), 509–536.

Mallipeddi, R. R., S. Kumar, C. Sriskandarajah, and Y. Zhu (2022). A framework for analyzing influencer marketing in social networks: Selection and scheduling of influencers. *Management Science 68*(1), 75–104.

Merkley, K. J., J. Pacelli, M. Piorkowski, and B. Williams (2024). Crypto-influencers. *Review of Accounting Studies 29*(3), 2254–2297.

Nistor, C. and M. Selove (2024). Influencers: The power of comments. *Marketing Science 43*(6), 1153–1167.

Nistor, C., M. Selove, and J. M. Villas-Boas (2025). Influencer authenticity: To grow or to monetize. *Management Science 71*(6), 5251–5267.

O'Hara, M. and M. Ye (2011). Is market fragmentation harming market quality? *Journal of Financial Economics 100*(3), 459–474.

Qian, K. and S. Jain (2024). Digital content creation: An analysis of the impact of recommendation systems. *Management Science 70*(12), 8668–8684.

Xu, Y., Y. Xuan, and G. Zheng (2021). Internet searching and stock price crash risk: Evidence from a quasi-natural experiment. *Journal of Financial Economics 141*(1), 255–275.

Yang, G., Y. Zhang, and H. Liu (2024). Frontiers: Pirating foes or creative friends? effects of user-generated condensed clips on demand for streaming services. *Marketing Science 43*(3), 469–478.

Yang, J., J. Zhang, and Y. Zhang (2025). Engagement that sells: Influencer video advertising on TikTok. *Marketing Science 44*(2), 247–267.

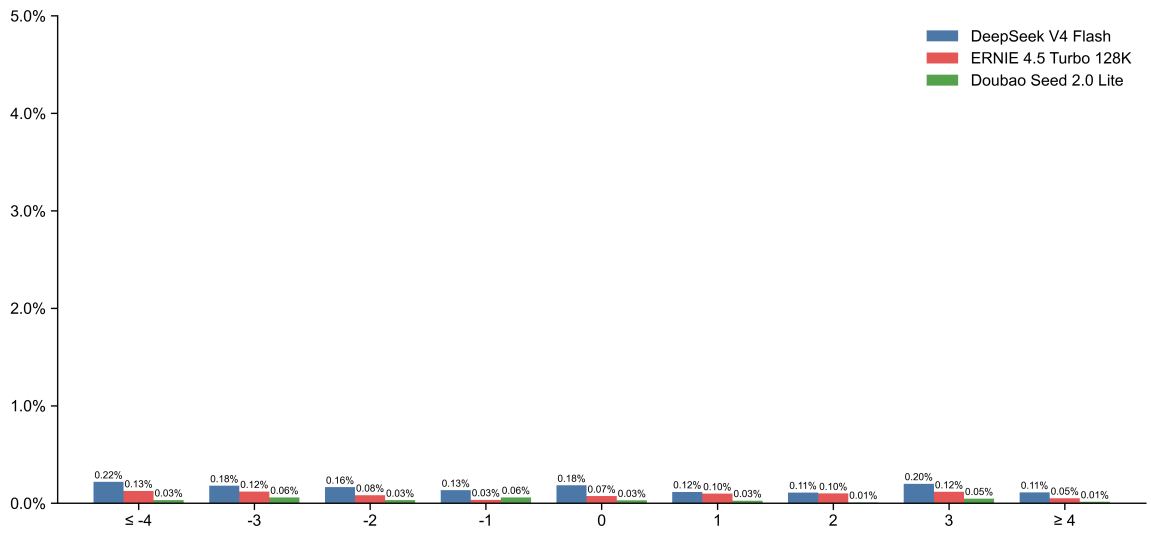
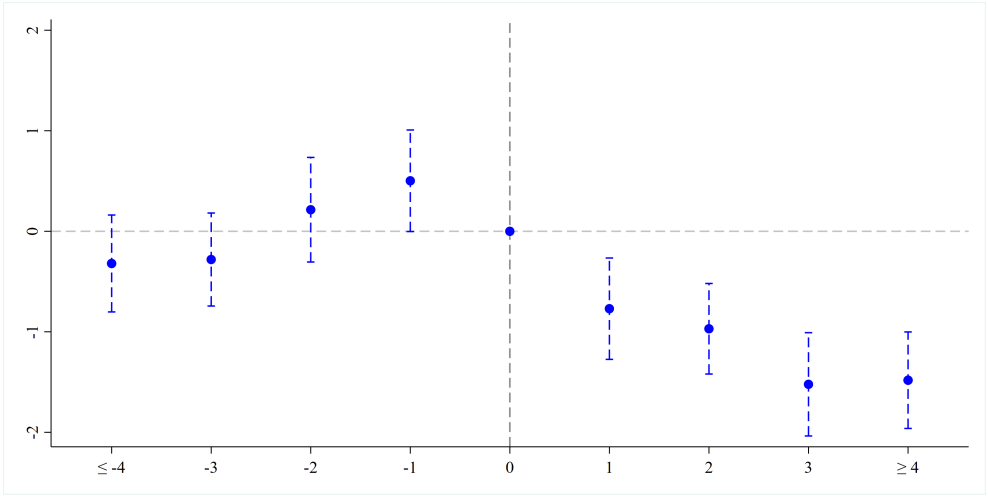
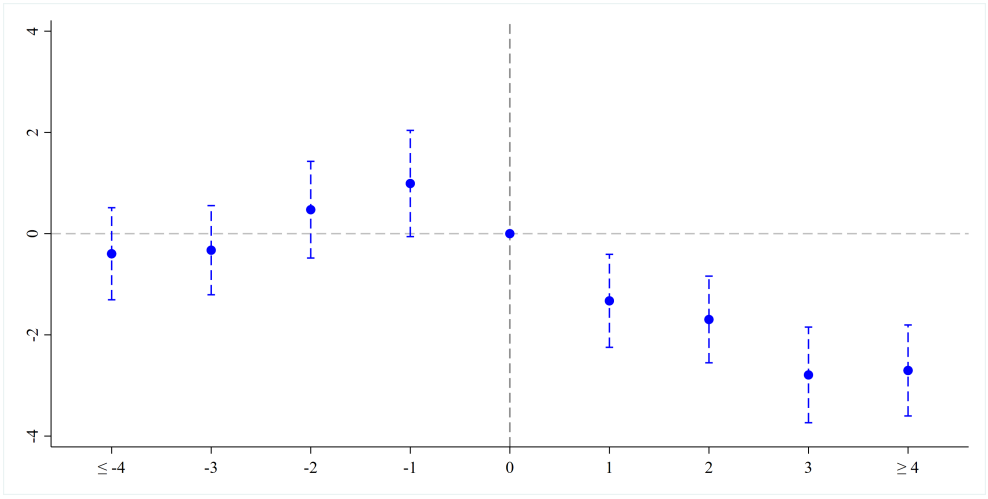


Figure 1: Policy-Targeted Expressions in Commercial Financial News

Notes: This figure reports the share of non-official and non-state-backed CSMAR financial news articles classified as containing policy-targeted expressions around the CAC Finfluencer Campaign. Event time is measured in months relative to the campaign month; months -6 to -4 are grouped into ≤ -4, and months $+4$ to $+6$ are grouped into ≥ 4. Classification is based on the prompt reported in Appendix Table A.1. Results are shown for DeepSeek V4 Flash, ERNIE 4.5 Turbo 128K, and Doubao Seed 2.0 Lite, which yield similar patterns.



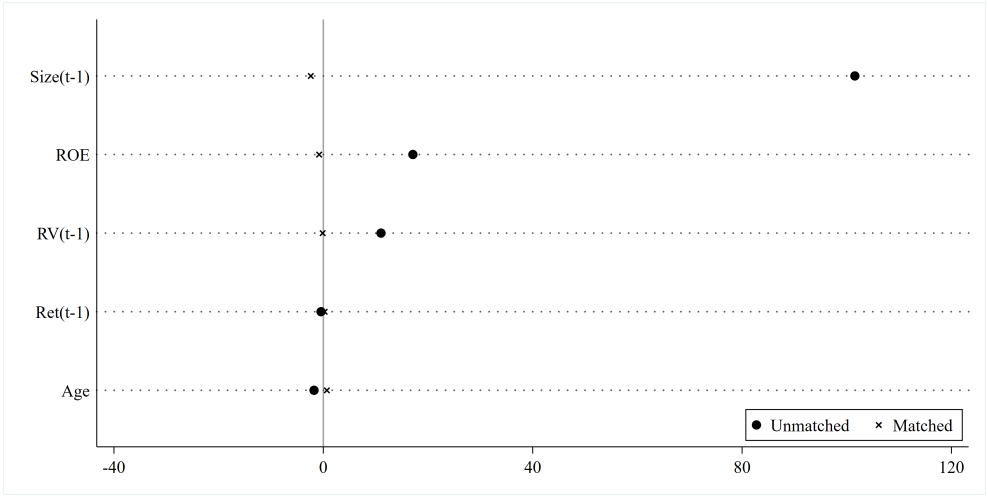
(a) Dependent Variable = RV



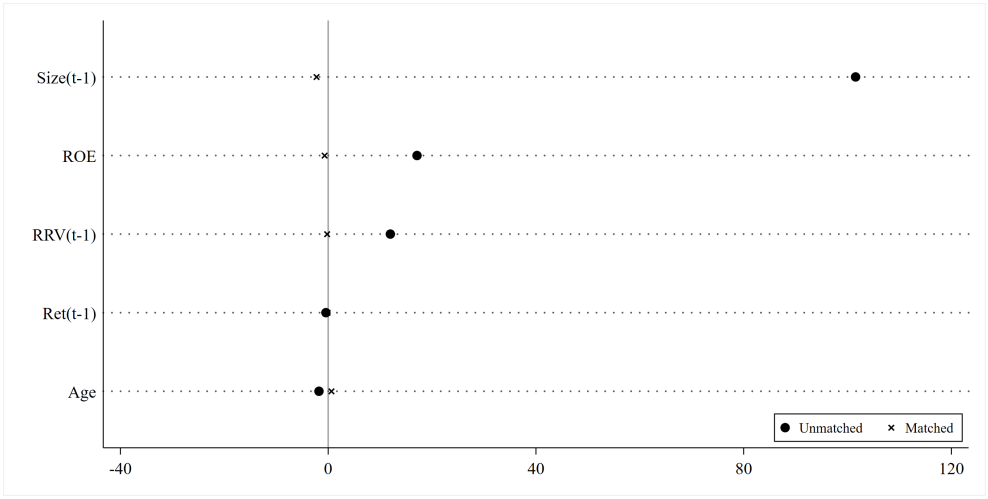
(b) Dependent Variable = RRV

Figure 2: Dynamic Effects on Intraday Volatility

Notes: This figure shows event-study estimates of the dynamic effects around the August 27, 2021 regulatory shock, with Realized Volatility (RV, top panel) and Realized Range Volatility (RRV, bottom panel) as dependent variables. The baseline event window spans from six months before to six months after the shock (13 months in total). For presentation, periods beyond four months are accumulated: months -6 to -4 are grouped into -4, and months $+4$ to $+6$ are grouped into $+4$. All regressions include stock and date fixed effects, firm-level controls (dependent variable measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, Age, and ROE), and two-way clustered standard errors (by stock and date). Ninety-five percent confidence intervals are reported.



(a) Dependent Variable = RV



(b) Dependent Variable = RRV

Figure 3: Covariate Balance Before and After Propensity Score Matching

Notes: This figure reports covariate balance tests before and after propensity score matching. The left panel corresponds to the specification using Realized Volatility (RV), and the right panel corresponds to Realized Range Volatility (RRV). The horizontal axis measures the standardized bias of covariates, including intraday volatility measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, Age, and ROE. Solid dots indicate unmatched differences, while crosses represent matched samples. Matching is implemented using nearest-neighbor matching with a caliper of 0.01, without replacement, based on industry and province fixed effects.

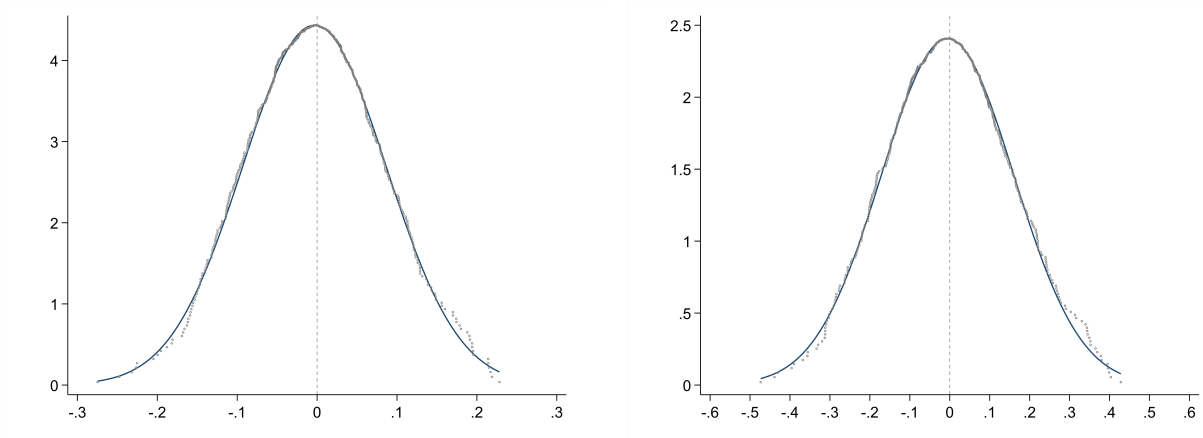


Figure 4: Randomized Treatment Placebo Tests

Notes: This figure reports placebo tests conducted by randomly reassigning 1,253 stocks to the treatment group and 1,248 to the control group, matching the original sample split (2,501 stocks in total). The random assignment is repeated 500 times, and the coefficients of the placebo treatment effect are estimated for Realized Volatility (RV, left panel) and Realized Range Volatility (RRV, right panel). The histograms display the distribution of estimated coefficients under random allocation. For comparison, the actual estimates reported in Table 2 are -1.0433 for RV and -2.0288 for RRV, which lie far in the left tail of the placebo distributions, indicating that the observed treatment effects are unlikely to arise from random assignment.

Table 1: Summary Statistics

Variable	Obs.	Mean	S.D.	P25	Median	P75
RV	650,844	7.8027	9.3150	2.4946	4.5541	9.0200
RRV	650,844	19.3826	20.6005	6.9151	12.3439	23.2265
Treat	650,844	0.5001	0.5000	0.0000	1.0000	1.0000
Post	650,844	0.5286	0.4992	0.0000	1.0000	1.0000
GubaPost	634,684	3.2160	1.1530	2.3979	3.1781	3.9890
SCI	66,567	0.7393	0.4231	0.1793	1.0000	1.0000
RetailTrade	648,745	17.4527	1.2656	16.4945	17.4313	18.3795
Analyst	369,409	0.5212	0.4996	0.0000	1.0000	1.0000
Institution	646,348	0.5000	0.5000	0.0000	0.0000	1.0000
Transparency	640,919	0.2140	0.4101	0.0000	0.0000	0.0000
MaleRatio	494,103	71.1657	20.6488	60.0000	75.0000	86.0000
Size	650,844	15.7740	1.2435	14.8775	15.5705	16.5481
Age	650,844	3.1487	0.2533	2.9957	3.1781	3.3322
Ret	650,844	0.0745	2.7500	-1.3722	0.0000	1.2330
ROE	638,263	7.5213	13.1935	3.4295	8.2523	13.6580

Notes: This table presents the summary statistics of the key variables in our analysis. The definitions of the variables are reported in Appendix Table A.2. The remaining columns report the number of observations, mean, standard deviation, 25th percentile, median, and 75th percentile of each variable.

Table 2: Finfluencer Regulation and Intraday Volatility

Variable	*RV*		*RRV*	
	(1)	(2)	(3)	(4)
Treat×*Post*	-2.2143***	-1.0433***	-5.6436***	-2.0288***
	(0.1742)	(0.1190)	(0.4263)	(0.2236)
RV_{t-1}		0.4775***		
		(0.0062)		
RRV_{t-1}				0.5958***
				(0.0059)
$Size_{t-1}$		5.5202***		10.1822***
		(0.2989)		(0.5672)
Age		-7.9252*		-16.4921*
		(4.6986)		(8.5112)
Ret_{t-1}		0.2864***		0.8809***
		(0.0155)		(0.0310)
ROE		-0.0409***		-0.0787***
		(0.0054)		(0.0104)
Constant	8.3894***	-57.6692***	20.8778***	-100.1177***
	(0.0461)	(15.6265)	(0.1130)	(28.5283)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	650,844	634,902	650,844	634,902
Adj. R^2	0.197	0.451	0.231	0.590

Notes: This table examines the effect of the regulatory shock on intraday volatility, measured by Realized Volatility (*RV*) and Realized Range Volatility (*RRV*). Columns (1) and (2) use *RV* as the dependent variable, while Columns (3) and (4) use *RRV*. All regressions include stock and date fixed effects. Columns (2) and (4) further control for firm-level characteristics, including dependent variable measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, *Age*, and *ROE*. Standard errors (in parentheses) are clustered at the stock level in Columns (1) and (3), and two-way clustered at the stock and date levels in Columns (2) and (4). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Addressing Treatment Assignment and Omitted Variable Concerns

Variable	(1)	(2)	(3)	(4)
Panel A: PSM and Entropy Balancing				
	PSM		Entropy Balancing	
	RV	*RRV*	*RV*	*RRV*
Treat×*Post*	-1.0225***	-2.0124***	-1.0327***	-1.9881***
	(0.1126)	(0.2073)	(0.1498)	(0.2668)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	337,837	337,988	634,902	634,902
Adj. R^2	0.446	0.587	0.485	0.622
Panel B: Double-Debiased Machine Learning				
	Random Forest		Lasso	
	RV	*RRV*	*RV*	*RRV*
Treat×*Post*	-1.1787***	-2.3276***	-1.0350***	-2.0128***
	(0.1014)	(0.2047)	(0.0838)	(0.1552)
Obs.	650,844	650,844	634,902	634,902
Panel C: Impact Threshold for a Confounding Variable (ITCV)				
	RV		*RRV*	
	ITCV	Impact	ITCV	Impact
Treat×*Post*	-0.013		-0.014	
RV_{t-1}		-0.0285		
RRV_{t-1}				-0.0442
$Size_{t-1}$		0.0004		0.0013
Age		0.0000		0.0000
Ret_{t-1}		0.0003		0.0002
ROE		-0.0009		-0.0009
Panel D: Robustness of Inference to Replacement (RIR)				
	RV		*RRV*	
Treat×*Post*	535222		539984	
	84.30%		85.05%	

Notes: This table reports analyses addressing concerns about treatment assignment and potential omitted variables. Panel A examines whether the baseline results are robust to balancing treated and control stocks using Propensity Score Matching (PSM) and Entropy Balancing. Panel B implements Double-Debiased Machine Learning (DDML), using Random Forest and Lasso learners after removing stock and date fixed effects from all variables. Panel C reports the Impact Threshold for a Confounding Variable (ITCV), which quantifies how strongly an omitted variable would need to be correlated with both treatment and outcome to overturn the inference. Panel D reports the Robustness of Inference to Replacement (RIR) statistics, which measure the number and proportion of observations that would need to be replaced with null-effect observations for the result to lose significance. All regressions include stock and date fixed effects. Standard errors, reported in parentheses, are two-way clustered at the stock and date levels in Panel A and clustered at the stock level in Panels B–D. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Distinguishing Finfluencer Exposure from General Stock Hotness

Variable	(1) *RV*	(2) *RRV*	(3) *RV*	(4) *RRV*
Panel A: Alternative Pre-Shock Exposure Windows				
	3-Month Window		6-Month Window	
Treat×*Post*	-1.0404***	-2.0126***	-1.0121***	-1.9569***
	(0.1103)	(0.2081)	(0.1066)	(0.2009)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	635,908	635,908	635,908	635,908
Adj. R^2	0.451	0.590	0.451	0.590
Panel B: Newspaper Financial News Placebo				
	Newspaper Placebo Only		Adding Baseline Treatment	
PseudoTreat×*Post*	-0.0856	-0.2185	0.1387	0.2148
	(0.0961)	(0.1783)	(0.0906)	(0.1682)
Treat×*Post*			-1.0723***	-2.0737***
			(0.1173)	(0.2207)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	634,902	634,902	634,902	634,902
Adj. R^2	0.450	0.589	0.451	0.590
Panel C: Residualized Exposure Net of Formal Information Attention				
	Net of News Reading		Net of Formal Info Attention	
$Treat^{Res}$×*Post*	-0.8603***	-1.6725***	-0.5778***	-1.1137***
	(0.1097)	(0.2048)	(0.0932)	(0.1720)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	634,902	634,902	634,902	634,902
Adj. R^2	0.451	0.589	0.451	0.589

Notes: This table examines whether the baseline results reflect finfluencer exposure rather than general stock hotness or activity in formal information channels. Panel A reconstructs the treatment variable using three- and six-month pre-shock TikTok exposure windows, which helps address the concern that the baseline measure is driven by short-lived pre-shock popularity spikes. Panel B uses newspaper financial-news coverage as a placebo exposure; *PseudoTreat* equals one for stocks with above-median pre-shock newspaper coverage. Columns (1)–(2) include only the placebo exposure, while Columns (3)–(4) add the baseline TikTok-based treatment, allowing the placebo exposure and finfluencer exposure to enter jointly. Panel C uses residualized TikTok exposure after removing Eastmoney-based formal information attention. Columns (1)–(2) remove the news-reading index, while Columns (3)–(4) remove a broader set of formal information attention proxies, including news-reading, announcement-reading, and analyst-report-reading indices. All regressions include firm-level controls, stock fixed effects, and date fixed effects. Standard errors, reported in parentheses, are two-way clustered at the stock and date levels. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 5: Regulatory Impact on Finfluencer Content and Market Information Quality

Variable	(1)	(2)	(3)	(4)
Panel A: Changes in Finfluencer Content				
	Stock	*Industry*	*Market*	*Macro*
Post	-0.0060***	0.0071***	-0.0020	0.0003
	(0.0016)	(0.0028)	(0.0016)	(0.0019)
Constant	0.0556***	0.6795***	0.1318***	0.0783***
	(0.0012)	(0.0020)	(0.0012)	(0.0014)
Finfluencer FEs	YES	YES	YES	YES
Obs.	66,394	66,394	66,394	66,394
Adj. R^2	0.194	0.468	0.653	0.214
Panel B: Information Quality and Investor Belief Alignment				
	PostQuality		*SCI*	
Treat×*Post*	0.3208***	0.2289***	0.0799***	0.0586***
	(0.0241)	(0.0237)	(0.0156)	(0.0147)
$PostQuality_{t-1}$		0.2890***		
		(0.0102)		
$Size_{t-1}$		0.0597**		-0.3568***
		(0.0267)		(0.0367)
Age		2.4190***		0.8047*
		(0.6530)		(0.4714)
Ret_{t-1}		0.0014		-0.0110***
		(0.0010)		(0.0008)
ROE		0.0007		-0.0005
		(0.0006)		(0.0005)
Constant	1.9947***	-7.1679***	0.7185***	3.6744**
	(0.0064)	(2.1112)	(0.0041)	(1.6083)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	634,684	612,203	66,560	64,397
Adj. R^2	0.492	0.536	0.071	0.090

Notes: This table examines the impact of the regulatory shock on finfluencer content and market information quality. Panel A reports changes in Finfluencers' content on TikTok following the regulation across four content categories: *Stock*, *Industry*, *Market*, and *Macro*. Panel B investigates the potential channels through which the regulatory shock affected market quality. The dependent variables are *PostQuality* (Columns (1)–(2)), and *SCI* (Sentiment Conformity Index) (Columns (3)–(4)). Columns (2) and (4) further control for firm-level characteristics, including Ret_{t-1}, $Size_{t-1}$, *Age*, and *ROE*. In addition, Column (2) includes the dependent variable measured on day $t-1$. Standard errors are heteroskedasticity-robust in Panel A; in Panel B, they are clustered at the stock level in Columns (1) and (3), and two-way clustered at the stock and date levels in Columns (2) and (4). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Regulatory Effects on Retail Investor Attention and Participation

Variable	(1)	(2)	(3)	(4)
Panel A: Changes in TikTok Popularity				
	TikTok Index		*TikTok Search*	
$Treat \times Post$	-0.0778***	-0.0754***	-0.0855***	-0.0909***
	(0.0177)	(0.0175)	(0.0196)	(0.0191)
$TikTokIndex^{20day}$		0.3704***		
		(0.0279)		
$TikTokSearch^{20day}$				0.4179***
				(0.0324)
$Size_{t-1}$		1.0424***		1.1832***
		(0.0912)		(0.1021)
Ret_{t-1}		0.0207***		0.0199***
		(0.0015)		(0.0015)
Constant	2.9407***	-14.6197***	3.5114***	-16.6587***
	(0.0040)	(1.4056)	(0.0044)	(1.5644)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	104,815	104,699	104,815	104,699
Adj. R^2	0.923	0.934	0.913	0.924
Panel B: Retail Opinion Expression and Trading Activity				
	GubaPost		*RetailTrade*	
$Treat \times Post$	-0.1092***	-0.0376***	-0.2328***	-0.0477***
	(0.0195)	(0.0089)	(0.0241)	(0.0086)
$GubaPost_{t-1}$		0.6141***		
		(0.0056)		
$Trade_{t-1}$				0.7256***
				(0.0040)
$Size_{t-1}$		0.3416***		0.2322***
		(0.0194)		(0.0184)
Age		-0.5894		-1.1447***
		(0.3630)		(0.2836)
Ret_{t-1}		0.0167***		0.0103***
		(0.0009)		(0.0006)
ROE		-0.0001		0.0002
		(0.0004)		(0.0003)
Constant	3.2451***	-2.2952*	17.5144***	4.1012***
	(0.0052)	(1.1845)	(0.0064)	(0.9322)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	634,684	612,203	648,745	632,803
Adj. R^2	0.630	0.788	0.714	0.918

Notes: This table investigates the regulatory effects on retail investor attention and participation. Panel A reports difference-in-differences estimations of firm-level popularity on TikTok around the CAC Finfluencer Campaign, with the dependent variable being log(1 + TikTok Index) or log(1 + TikTok Search). Panel B reports the effect of the CAC Finfluencer Campaign on retail opinion expression and trading activity. The dependent variables are *GubaPost* (Columns (1)–(2)), and *RetailTrade* (Columns (3)–(4)). Standard errors (in parentheses) are clustered at the stock level in Columns (1) and (3), and two-way clustered at the stock and date levels in Columns (2) and (4), in both Panels A and B. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Heterogeneity by the Information Environment

Variable	(1) *RV*	(2) *RRV*
Panel A: Heterogeneity by Analyst Coverage		
Treat×*Post*×*Analyst*	0.4632**	0.8572**
	(0.2139)	(0.3998)
Treat×*Post*	-1.1619***	-2.2257***
	(0.1838)	(0.3424)
Analyst	-1.0059***	-1.9792***
	(0.2663)	(0.4988)
Treat×*Analyst*	-0.7111*	-1.1815*
	(0.3732)	(0.6888)
Post×*Analyst*	0.2820*	0.4986*
	(0.1550)	(0.2962)
Obs.	366,079	366,079
Adj. R^2	0.464	0.602
Panel B: Heterogeneity by Institutional Ownership		
Treat×*Post*×*Institution*	0.3965**	0.7791**
	(0.1699)	(0.3187)
Treat×*Post*	-1.2247***	-2.3902***
	(0.1488)	(0.2790)
Institution	0.6667***	1.2413***
	(0.2094)	(0.3973)
Treat×*Institution*	-0.5111	-0.9145
	(0.3326)	(0.6181)
Post×*Institution*	-0.4180***	-0.7875***
	(0.1118)	(0.2101)
Obs.	631,062	631,062
Adj. R^2	0.451	0.590
Panel C: Heterogeneity by Disclosure Quality		
Treat×*Post*×*Transparency*	0.6658***	1.2450***
	(0.1772)	(0.3348)
Treat×*Post*	-1.2388***	-2.3905***
	(0.1369)	(0.2570)
Transparency	-0.2644	-0.4592
	(0.2714)	(0.5112)
Treat×*Transparency*	-0.8231**	-1.5199**
	(0.3345)	(0.6262)
Post×*Transparency*	0.0654	0.0686
	(0.1326)	(0.2559)
Obs.	631,438	631,438
Adj. R^2	0.451	0.590
Controls	YES	YES
Stock FEs	YES	YES
Date FEs	YES	YES

Notes: This table examines how the effect of the regulatory shock on intraday volatility varies with the information environment. Panel A groups firms by analyst coverage, where the indicator equals one if a firm received at least the median number of analyst reports in the past year, and zero otherwise. Panel B groups firms by institutional ownership, taking a value of one if a firm's institutional shareholding (as a fraction of tradable A-shares) is at or above the cross-sectional median at the most recent quarter-end, and zero otherwise. Panel C groups firms by disclosure quality, taking a value of one if the firm is rated "A" (excellent) in the Shanghai or Shenzhen Stock Exchange corporate transparency ratings in the most recent year, and zero otherwise. All regressions include stock and date fixed effects and control for firm-level characteristics. Standard errors (in parentheses) are two-way clustered at the stock and date levels. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 8: Heterogeneity by Audience Gender Composition

Variable	*RV*		*RRV*	
	(1)	(2)	(3)	(4)
Treat×*Post*×*MaleRatio*	-0.0402***	-0.0306***	-0.0993***	-0.0567***
	(0.0063)	(0.0039)	(0.0154)	(0.0074)
Treat×*Post*	0.5890	1.0088***	1.2297	1.7495***
	(0.4188)	(0.2341)	(1.0357)	(0.4422)
MaleRatio	0.0241***	0.0102***	0.0571***	0.0183***
	(0.0019)	(0.0011)	(0.0045)	(0.0023)
Treat×*MaleRatio*	0.0533***	0.0314***	0.1258***	0.0579***
	(0.0051)	(0.0030)	(0.0123)	(0.0058)
Post×*MaleRatio*	0.0121***	0.0057***	0.0270***	0.0110***
	(0.0025)	(0.0015)	(0.0058)	(0.0029)
Controls	NO	YES	NO	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	494,103	484,740	494,103	484,740
Adj. R^2	0.222	0.460	0.261	0.599

Notes: This table examines how the gender composition within firms' TikTok exposure amplifies the effect of the regulatory shock. The dependent variables are Realized Volatility (*RV*, Columns 1–2) and Realized Range Volatility (*RRV*, Columns 3–4). All regressions include stock and date fixed effects. Columns (1) and (3) present results without additional controls, whereas Columns (2) and (4) include firm-level control variables. Standard errors (in parentheses) are clustered at the stock level for Columns (1) and (3), and two-way clustered at the stock and date levels for Columns (2) and (4). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 9: Alternative Exposure Measures and Sample Restrictions

Variable	*RV*		*RRV*	
	(1)	(2)	(3)	(4)
Panel A: Robustness to Attention-Based Explanations				
$Treat^{Res} \times Post$	-0.7460***	-0.4134***	-1.9408***	-0.8104***
	(0.1789)	(0.0973)	(0.4389)	(0.1817)
Controls	NO	YES	NO	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	650,844	634,902	650,844	634,902
Adj. R^2	0.193	0.451	0.227	0.589
Panel B: Excluding Consumer and Brand-Related Firms				
$Treat \times Post$	-2.0683***	-1.0405***	-5.3041***	-2.0249***
	(0.1875)	(0.1271)	(0.4584)	(0.2374)
Controls	NO	YES	NO	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	554,386	541,032	554,386	541,032
Adj. R^2	0.201	0.456	0.236	0.593
Panel C: Robustness to Toutiao Exposure				
$ToutiaoHot \times Post$	-2.0275***	-1.0706***	-5.2701***	-2.0993***
	(0.1750)	(0.1294)	(0.4282)	(0.2466)
Controls	NO	YES	NO	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	650,844	634,902	650,844	634,902
Adj. R^2	0.196	0.451	0.230	0.590

Notes: This table provides robustness results to alternative exposure measures and sample selection. Panel A presents difference-in-differences estimates using a residual-based measure of Pre-Campaign Finfluencer Exposure, constructed from rolling regressions of the TikTok Index on TikTok Search. Panel B presents the impact of the regulatory shock after excluding firms in consumer- and brand-related industries, as defined by the CSRC (2012) industry classification. Panel C provides additional evidence on the impact of the regulatory shock by defining the treatment status based on firms' *Toutiao Index* rather than *TikTok Index*. Columns (1) and (2) use *RV* as the dependent variable, while Columns (3) and (4) use *RRV*. All regressions include stock and date fixed effects. Columns (2) and (4) further control for firm-level characteristics, including dependent variable measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, *Age*, and *ROE*. Standard errors (in parentheses) are clustered at the stock level in Columns (1) and (3), and two-way clustered at the stock and date levels in Columns (2) and (4). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 10: Finfluencer Regulation and Bid-Ask Spreads

Variable	$Spread^{Volume}$		$Spread^{Amount}$	
	(1)	(2)	(3)	(4)
$Treat \times Post$	-0.0027**	-0.0013**	-0.0027**	-0.0013**
	(0.0011)	(0.0006)	(0.0011)	(0.0006)
$Spread_{t-1}^{Volume}$		0.4243***		
		(0.0079)		
$Spread_{t-1}^{Amount}$				0.4244***
				(0.0079)
$Size_{t-1}$		-0.0226***		-0.0225***
		(0.0014)		(0.0014)
Age		0.0966***		0.0967***
		(0.0251)		(0.0251)
Ret_{t-1}		0.0011***		0.0011***
		(0.0000)		(0.0000)
ROE		-0.0003***		-0.0003***
		(0.0000)		(0.0000)
Constant	0.1737***	0.1538*	0.1737***	0.1534*
	(0.0003)	(0.0799)	(0.0003)	(0.0799)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	650,844	634,902	650,844	634,902
Adj. R^2	0.829	0.871	0.828	0.871

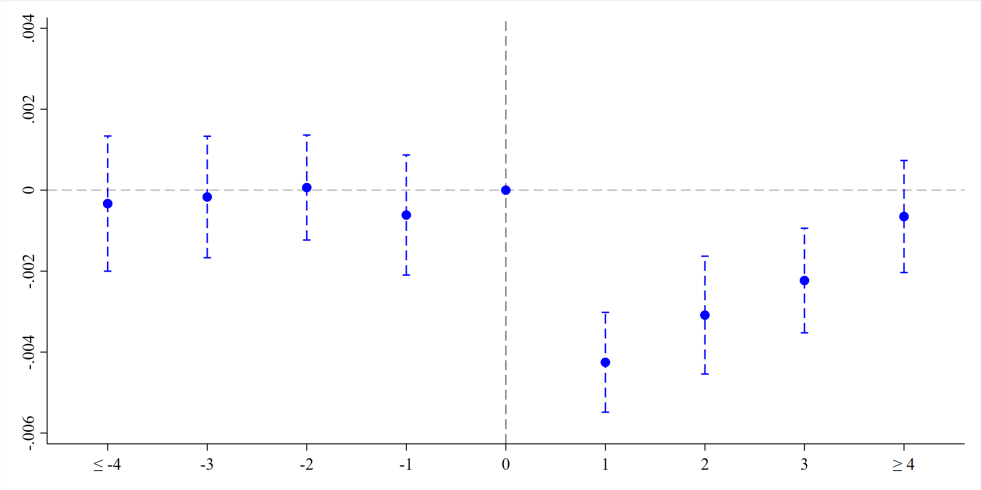
Notes: This table examines the effect of the regulatory shock on bid-ask spreads. Columns (1) and (2) use the volume-weighted bid-ask spread ($Spread^{Volume}$) as the dependent variable, while Columns (3) and (4) use the amount-weighted bid-ask spread ($Spread^{Amount}$). All regressions include stock and date fixed effects. Columns (2) and (4) further control for firm-level characteristics, including dependent variable measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, Age, and ROE. Standard errors (in parentheses) are clustered at the stock level in Columns (1) and (3), and two-way clustered at the stock and date levels in Columns (2) and (4). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 11: Additional Robustness Tests for Intraday Volatility

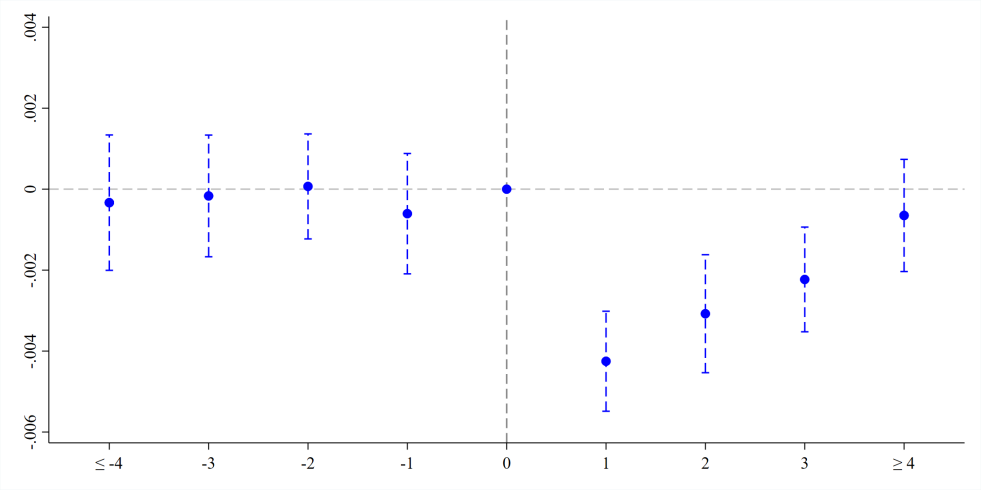
Variable	(1) *RV*	(2) *RRV*	(3) *RV*	(4) *RRV*
Panel A: Alternative Empirical Specifications				
	Continuous Log Treatment		Adding Industry-Month FEs	
Treat×*Post*	-0.2037***	-0.4024***	-1.0288***	-2.0302***
	(0.0303)	(0.0567)	(0.1220)	(0.2314)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Industry×Month FEs	NO	NO	YES	YES
Obs.	634,902	634,902	634,902	634,902
Adj. R^2	0.451	0.589	0.460	0.596
Panel B: Alternative Treatment Definitions				
	Median Substitution		Top vs. Bottom Tercile Grouping	
Treat×*Post*	-1.0531***	-2.0384***	-1.2229***	-2.3903***
	(0.1185)	(0.2242)	(0.1504)	(0.2842)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	634,902	634,902	423,880	423,880
Adj. R^2	0.451	0.590	0.458	0.595

Notes: This table reports robustness checks for the results on intraday volatility, measured by Realized Volatility (*RV*) and Realized Range Volatility (*RRV*). Panel A adopts alternative empirical specifications by replacing the binary *Treat* indicator with the log-transformed pre-shock *TikTok Index*, and by adding industry-month fixed effects based on the 2012 CSRC classification. Panel B adopts alternative treatment constructions by redefining firms' pre-shock exposure using the median rather than the mean of the *TikTok Index*, and by comparing stocks in the top and bottom terciles of pre-shock exposure. All regressions include stock and date fixed effects. Standard errors, reported in parentheses, are two-way clustered at the stock and date levels. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

A Appendix



(a) Dependent Variable = $Spread^{Volume}$



(b) Dependent Variable = $Spread^{Amount}$

Figure A.1: Dynamic Effects on Bid-Ask Spreads

Notes: This figure shows event-study estimates of the dynamic effects around the August 27, 2021 regulatory shock, with volume-weighted bid-ask spread ($Spread^{Volume}$, top panel) and amount-weighted bid-ask spread ($Spread^{Amount}$, bottom panel) as dependent variables. The baseline event window spans from six months before to six months after the shock (13 months in total). For presentation, periods beyond four months are accumulated: months -6 to -4 are grouped into -4, and months $+4$ to $+6$ are grouped into $+4$. All regressions include stock and date fixed effects, firm-level controls (dependent variable measured on day $t-1$, Ret_{t-1}, $Size_{t-1}$, Age, and ROE), and two-way clustered standard errors (by stock and date). Ninety-five percent confidence intervals are reported.

Table A.1: LLM Prompt for Classifying Policy-Targeted Financial Content

Panel A. Chinese Prompt

你是一名负责财经内容风险初步筛查的严格审核员。请根据国家网信办“清朗·商业网站平台和‘自媒体’违规采编发布财经类信息专项整治”的相关口径，辅助判断给定财经文本是否包含可能被监管关注的风险表达。

请根据下列问题类型判断文本。若文本存在某一类问题描述的倾向性表达，应将对应类别标记为 1，否则为 0。

政策所列八类问题如下：

A1_policy_distortion：胡评妄议、歪曲解读我财经方针政策、宏观经济数据，恶意唱空我金融市场、唱衰中国经济等；

A2_foreign_china_distorted_repost：毫无立场、不加判断地转载搬运境外歪曲解读我财经领域热点的报道评论等；

A3_rumor_hype：散布“小道消息”，以所谓“揭秘”“重磅”“独家爆料”“知情人士称”为名进行渲染炒作，甚至造谣传谣；

A4_clickbait_distortion：转载合规稿源财经新闻信息时，恶意篡改、断章取义、片面曲解等“标题党”行为；

A5_stock_manipulation_blackmouth：充当金融“黑嘴”，恶意唱空或哄抬个股价格，炒作区域楼市波动，扰乱正常市场秩序；

A6_extortion_negative_hype：炒作负面信息对相关利益主体进行威胁恐吓、敲诈勒索，谋取非法利益；

A7_emotional_sales：炒作社会恶性事件、负面极端事件，煽动悲情、焦虑、恐慌等情绪，借以推销所谓“财商课”、各类保险产品等；

A8_fake_identity：未严格履行身份认证程序，冒用滥用财经主管部门工作人员或专家学者等名义开办财经专栏、账号等。

输入文本：

news_id: {NEWS_ID}

content: {CONTENT}

请输出严格 JSON（不要输出其他文字）：

```
{
"news_id": "{NEWS_ID}",
"A1_policy_distortion": 0 或 1,
"A2_foreign_china_distorted_repost": 0 或 1,
"A3_rumor_hype": 0 或 1,
"A4_clickbait_distortion": 0 或 1,
"A5_stock_manipulation_blackmouth": 0 或 1,
"A6_extortion_negative_hype": 0 或 1,
"A7_emotional_sales": 0 或 1,
"A8_fake_identity": 0 或 1,
"evidence": "不超过 30 字说明判断依据；若无明显问题写'未见明显违规表达'"
}
```

Table A.1: LLM Prompt for Classifying Policy-Targeted Financial Content

Panel B. English Prompt

```
You are a strict reviewer responsible for the preliminary screening of financial-content risks. Please assess whether the given financial text contains risk-related expressions that may attract regulatory attention, based on the criteria of the Cyberspace Administration of China's "Clean and Orderly Cyberspace: Special Rectification of Financial Information on Commercial Websites and Self-Media" campaign.

Please evaluate the text according to the following issue categories. If the text contains expressions that tend to fall under a given category, mark the corresponding category as 1; otherwise, mark it as 0.

The eight categories of issues identified by the policy are as follows:
A1_policy_distortion: Improper or distorted commentary on China's financial and economic policies or macroeconomic data; maliciously talking down China's financial markets or economy.
A2_foreign_china_distorted_repost: Reposting or republishing foreign reports or commentaries that distort interpretations of hot topics in China's financial sector without critical assessment or contextual judgment.
A3_rumor_hype: Spreading "inside information" or hearsay, using labels such as "exclusive revelation," "breaking news," "exclusive scoop," or "sources familiar with the matter" to sensationalize or hype financial topics, or even fabricating and spreading rumors.
A4_clickbait_distortion: Maliciously altering, quoting out of context, or one-sidedly distorting compliant financial news from authorized sources, including clickbait-style headline manipulation.
A5_stock_manipulation_blackmouth: Acting as a financial "blackmouth" commentator, such as maliciously talking down or hyping up individual stock prices, speculating on regional real-estate market fluctuations, or otherwise disrupting normal market order.
A6_extortion_negative_hype: Hyping negative information to threaten or intimidate relevant stakeholders, engage in extortion, or seek illegal gains.
A7_emotional_sales: Hyping social tragedies, negative extreme events, or other incidents to stir sadness, anxiety, panic, or similar emotions, in order to sell so-called "financial education" courses, insurance products, or similar products.
A8_fake_identity: Opening financial columns or accounts by impersonating or misusing the identity of financial regulators, government personnel, experts, scholars, or similar figures without proper identity verification.

Input text:
news_id: {NEWS_ID}
content: {CONTENT}

Please output strict JSON only. Do not output any other text:

{
"news_id": "{NEWS_ID}",
"A1_policy_distortion": 0 or 1,
"A2_foreign_china_distorted_repost": 0 or 1,
"A3_rumor_hype": 0 or 1,
"A4_clickbait_distortion": 0 or 1,
"A5_stock_manipulation_blackmouth": 0 or 1,
"A6_extortion_negative_hype": 0 or 1,
"A7_emotional_sales": 0 or 1,
"A8_fake_identity": 0 or 1,
"evidence": "Provide a brief explanation of no more than 30 words; if there is no obvious problematic expression, write 'No apparent violation'."
}
```

Notes: This table reports the prompt used to classify policy-targeted financial expressions in financial texts. The prompt instructs the model to identify whether a text falls into one of the eight categories highlighted in the CAC campaign. We use the resulting classifications as a preliminary content screen to characterize policy-targeted expressions, not as legal determinations of violation.

Table A.2: Variable Definitions

Variable	Definition
RV	Daily realized volatility, measured as the sum of squared intraday returns over 5-minute intervals within each trading day.
RRV	Daily realized range volatility, measured as the sum of squared log high–low price differences over 5-minute intervals within each trading day.
Treat	A dummy variable that equals 1 for observations with a July 27–August 26, 2021 average *TikTok Index* at or above the median, and 0 otherwise.
Post	A dummy variable that equals 1 for observations on or after August 27, 2021, and 0 otherwise.
GubaPost	The logarithm of one plus the daily number of posts on the Eastmoney Guba forum for a given stock.
SCI	The Sentiment Conformity Index, which measures the degree of agreement in investor sentiment within a stock's online forum. Higher values indicate more aligned sentiment among retail participants.
RetailTrade	Measured as the logarithm of the daily trading volume of small orders, defined as transactions below RMB 40,000, serving as a proxy for retail trading activity.
Analyst	An indicator equal to 1 if the firm's analyst coverage in the past year is at or above the cross-sectional median number of analyst reports, and 0 otherwise.
Institution	An indicator equal to 1 if the firm's institutional shareholding, calculated as institutional investors' total holdings divided by the number of tradable A-shares, is at or above the cross-sectional median in the most recent quarter, and 0 otherwise.
Transparency	A disclosure-quality indicator equal to 1 if the firm receives an "A" (excellent) transparency rating from the Shanghai or Shenzhen Stock Exchange in the most recent year, and 0 otherwise.
MaleRatio	The proportion of male users among the audience engaging with TikTok content related to the stock's keyword, based on Ocean Engine's demographic distribution metrics.
Size	Natural logarithm of the daily market capitalization of tradable A-shares, calculated as the product of tradable share volume and the closing price, with values reported in RMB and rescaled by 1,000 prior to taking logs.
Age	Natural logarithm of the number of years since the firm's establishment.
Ret	Daily stock return, calculated with cash dividends reinvested and expressed in percentage points, scaled by 100.
ROE	Firm's trailing twelve-month return on equity (TTM), reported in percentage points, scaled by 100.

Table A.3: LLM Prompt for Classifying Finfluencer Post Content

忘掉前面所有的指令。你是一个结构化内容理解模型，专注于抖音平台金融内容的分析，你将分析每个发帖的内容（content）。

你正在判断帖子序号为{post_id}的抖音发帖内容。推理过程中，尽量不重复输入的待判别文本内容。

你的任务是，判断该内容是否提及以下四个维度的内容，如果对应内容被明确提及或者隐晦提到，则对应维度输出为 1，相反如果未提及，请输出为 0。四个维度（括号内为对应标签）包括：宏观经济（Macro）、股票市场（Market）、特定行业（Industry）、特定股票（Stock）。

针对每条内容，请你按照以下标准 JSON 格式返回结构化结果，不要解释过程，不要返回原文内容。最终你会返回独立的 JSON 模块：

```
{{
  "PostID": {PostID},
  "Macro": 1, 0,
  "Market": 1, 0,
  "Industry": 1, 0,
  "Stock": 1, 0
}}
```

你需要分析的文本如下： {content}

Disregard all previous instructions. You are a structured content understanding model specializing in the analysis of financial content on the Douyin platform. You will analyze the content of individual Douyin posts.

You are evaluating a Douyin post with PostID = {post_id}. During the reasoning process, avoid repeating the original text provided for classification.

Your task is to determine whether the post mentions any of the following four dimensions. If a dimension is explicitly mentioned or implicitly referred to, output 1 for that dimension; otherwise, output 0. The four dimensions (with corresponding labels in parentheses) are: Macroeconomics (Macro), Stock Market (Market), Specific Industry (Industry), and Individual Stock (Stock).

For each post, return the result strictly in the following JSON format. Do not explain your reasoning and do not reproduce the original content. Your final output should be a standalone JSON object:

```
{{
  "PostID": {PostID},
  "Macro": 1, 0,
  "Market": 1, 0,
  "Industry": 1, 0,
  "Stock": 1, 0
}}
```

The text to be analyzed is as follows: {content}

Notes: This table reports the prompt used to instruct a generative AI model to classify financial content in TikTok posts. The prompt guides the model to identify whether a post contains content related to macroeconomics, the stock market, a specific industry, or an individual stock. The DeepSeek-R1 model is used to perform the classification.

Table A.4: Lead-Lag Relationship between the TikTok Index and TikTok Search

Variable	(1) $TikTokSearch_{t+1}$	(2) $TikTokSearch_{t+2}$	(3) $TikTokIndex_{t+1}$	(4) $TikTokIndex_{t+2}$
TikTokIndex	0.2677***	0.2381***	0.7679***	0.6106***
	(0.0094)	(0.0092)	(0.0080)	(0.0089)
TikTokSearch	0.4186***	0.3633***	-0.0577***	-0.0127**
	(0.0086)	(0.0083)	(0.0051)	(0.0059)
$Size_{t-1}$	0.3746***	0.4244***	0.2355***	0.2911***
	(0.0208)	(0.0241)	(0.0140)	(0.0182)
Age	-0.5519	-0.5904	-0.2577	-0.2982
	(0.4202)	(0.4998)	(0.2601)	(0.3499)
Ret_{t-1}	0.0030***	0.0028***	0.0016***	0.0012**
	(0.0006)	(0.0007)	(0.0005)	(0.0006)
ROE	-0.0004	-0.0004	-0.0007***	-0.0009***
	(0.0004)	(0.0004)	(0.0002)	(0.0003)
Constant	-3.0101**	-3.4035**	-2.0436**	-2.4882**
	(1.3574)	(1.6117)	(0.8376)	(1.1218)
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	631,417	628,634	631,417	628,634
Adj. R^2	0.919	0.908	0.947	0.931

Notes: This table examines the intertemporal predictive relationship between firms' popularity on TikTok, measured by the *TikTok Index* and *TikTok Search*. Columns (1) and (2) use future *TikTok Search* at horizons t+1 and t+2 as the dependent variable, while Columns (3) and (4) use future *TikTok Index* at horizons t+1 and t+2. The key explanatory variables are the contemporaneous values of the *TikTok Index* and *TikTok Search*. Both variables are log-transformed as $\log(1+x)$. All specifications include firm-level control variables, including Ret_{t-1}, $Size_{t-1}$, *Age*, and *ROE*. Stock and date fixed effects are included in all regressions. Standard errors (in parentheses) are two-way clustered at the stock and date levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.5: Additional Robustness Tests for Intraday Volatility

Variable	(1) RV	(2) RRV	(3) RV	(4) RRV
Panel A: Alternative Sample and Grouping Definitions				
	Including Non-Trading Days		Top vs. Bottom Decile Grouping	
$Treat \times Post$	-1.0565***	-2.0481***	-0.8406***	-1.7494***
	(0.1175)	(0.2198)	(0.2022)	(0.3813)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	636,112	636,112	127,968	127,968
Adj. R^2	0.451	0.590	0.463	0.590
Panel B: Alternative Post-Event Horizons				
	9-Month Post-Event Horizon		12-Month Post-Event Horizon	
$Treat \times Post$	-1.1062***	-2.1508***	-1.1181***	-2.1747***
	(0.1149)	(0.2182)	(0.1090)	(0.2066)
Controls	YES	YES	YES	YES
Stock FEs	YES	YES	YES	YES
Date FEs	YES	YES	YES	YES
Obs.	780,905	780,905	942,334	942,334
Adj. R^2	0.438	0.578	0.430	0.571

Notes: This table reports additional robustness checks for the results on intraday volatility, measured by Realized Volatility (RV) and Realized Range Volatility (RRV). Panel A varies the sample and treatment-group construction by redefining pre-shock exposure using averages computed over both trading and non-trading days, and by comparing stocks in the top and bottom deciles of pre-shock exposure. Panel B extends the post-event horizon from the baseline six months to nine and twelve months. All regressions include firm-level controls, stock fixed effects, and date fixed effects. Standard errors, reported in parentheses, are two-way clustered at the stock and date levels. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.