Data Science in Finance and Economics

DSFE, 6(4): 595–621.
DOI: 10.3934/DSFE.2026021
Received: 12 March 2026
Revised: 04 June 2026
Accepted: 17 July 2026
Published: 05 November 2026

https://www.aimspress.com/journal/dsfe

Research article

Does investor interaction on social media influence financial market volatility? A text-based network analysis of CSI 300 constituents

Zhongxiu Chen[1] and Zhenghui Li[1,2]*

[1] School of Economics and Statistics, Guangzhou University, Guangzhou,510000, China
[2] Guangzhou Institute of International Finance, Guangzhou University, Guangzhou, 510000, China

* **Correspondence:** Email: lizh@gzhu.edu.cn.

Abstract: The advancement of information technology and the widespread adoption of social media have accelerated information exchange among investors. This study investigated the impact of investor-driven information flow on financial market volatility. We collected 5,258,185 comments from the Eastmoney Stock Forum covering the CSI 300 constituent stocks between June 1, 2023, and May 31, 2024, and constructed 48 weekly investor-driven information flow networks. Using panel data comprising 14,400 firm-week observations and regression analysis, we found that market volatility responds significantly to the structural properties of these networks. Firms with higher network centrality experience a 1.3% increase in return volatility and a 17.1% rise in trading volume. While volatility declines over time (−8.4%), trading activity remains higher (34.5%). We further confirmed that network centrality drives greater message activity and promotes opinion convergence, supporting two mechanisms: prompt information incorporation and consensus formation. These insights contribute to a deeper understanding of the mechanisms through which information networks on social media shape market dynamics.

Keywords: investor interaction; information flow network; social media; financial market volatility; trading behavior

JEL Codes: G14, G12, D85

1. Introduction

In early 2025, certain technology stocks experienced rapid price increases driven primarily by intense discussion and attention on social media platforms, such as the Eastmoney Stock Forum and Xue Qiu, rather than by fundamental news. Prior to abnormal price movements, stocks often witness a surge in online discussion, where comments, reposts, and hashtags collectively form emergent information flow networks. In Figure 1, we present an example page on an interaction platform. This observation raises a fundamental question: can frequent investor interactions on social media generate structured information flow networks that systematically influence market behavior, including volatility and trading activity?



Figure 1. Example page on the East Money Stock Forum. (Source: https://guba.eastmoney.com/list,000002.html.)

Growing evidence suggests that information exchange among investors is not merely a form of sentiment-driven expression. It may, in fact, constitute a structured information diffusion channel that influences market pricing mechanisms (Du et al., 2023; Wang et al., 2025; Tuan et al., 2024; An et al., 2024). This effect is particularly pronounced in China's A-share market, where retail participation is high, and information spreads rapidly through peer-to-peer interaction. While prior research has explored the influence of various information sources—such as news media, analyst reports, and online sentiment—on financial markets, much of it focuses on isolated firms or specific media channels. Fewer studies address how investor interactions themselves form a transmission network and how such

networks impact market outcomes at a systemic level (Zhang et al., 2024; Yu et al., 2024; Wang et al., 2024; Han and Yang, 2013; Luo, Zhao, and Zhou, 2023; Mikhaylov et al., 2023).

This study contends that investors on social media operate not in isolation but as nodes within a dynamic information flow network. Through commenting, reposting, and engagement, they shape the pathways of information dissemination and the allocation of market attention. We investigate how investor-driven information networks affect two key market behaviors: stock return volatility and trading intensity. Specifically, we assess whether firms occupying central positions within these networks exhibit distinct market dynamics due to their increased exposure to investor attention and information flows. We develop hypotheses grounded in two potential channels: (1) the accelerated information diffusion mechanism, whereby centrally positioned firms are more frequently discussed, enabling faster information transmission and immediate price incorporation; and (2) the belief-alignment mechanism, in which frequent investor interactions reduce information asymmetry and foster consensus formation, thereby influencing both volatility and liquidity.

Our empirical design incorporates idiosyncratic return volatility (IRV) to measure firm-specific risk, isolating the residual component of return variability net of market-wide influences. Trading behavior is assessed via standardized trading volume and turnover ratio, proxies for investor engagement and attention. We further integrate both global network characteristics (e.g., density, hierarchy, efficiency) and local centrality metrics (e.g., degree, closeness, betweenness) into our regression framework.

This study offers both theoretical and empirical contributions. Theoretically, it extends existing literature on investor attention and information diffusion by incorporating network-based interaction structures into the analysis of market dynamics. This network perspective complements traditional sentiment-based models and sheds light on how investor-driven information flow can systematically affect price formation and trading behavior. Empirically, it demonstrates how investor networks formed through social media interactions reshape market attention and influence market volatility. By leveraging real-time comment data, the study provides new evidence on the behavioral mechanisms underlying market volatility and trading activity.

The subsequent arrangement of this article is as follows: Section 2 presents the literature review and Section 3 provides the research hypotheses. Section 4 introduces methodology. Section 5 presents the research findings. Section 6 discusses the theoretical and practical implications and concludes with some future research directions.

2. Literature review

In financial market research, the question of how information influences prices has long been a central theme. With the widespread adoption of social media, the mode of information transmission among investors has changed significantly—information is no longer disseminated solely through official announcements or professional media platforms. This phenomenon has given rise to a new research direction: Is the flow of information within social networks systematically influencing market behavior?

Existing literature can be broadly categorized into three streams. The first is investor attention literature. This stream examines how investor attention affects financial market dynamics. Attention is typically measured using online behavioral proxies, such as social media discussion intensity, search volume, and post frequency. Warkulat and Pelster (2024) used r/wallstreetbets discussion activity and individual trading data to show that social media attention stimulates trading and risk-taking. Zhang

and Luo (2024) measured investor attention using internet search volume in the Chinese stock market and found that higher attention affects analysts' forecast errors and dispersion. Ayala et al. (2024) reviewed studies using Google Search Volume Index and post frequency to show that investor attention help explain trading volume, volatility, and return predictability. These works capture attention-driven behavior in markets. The second stream are information diffusion models. This strand focuses on how information spreads among investors, firms, or media channels using network-based frameworks. Hirshleifer, Peng, and Wang (2025) constructed an information diffusion model where firms are nodes in an investor communication network, using social network data. Ying (2024) built a gradual diffusion model based on common institutional ownership, representing firms as nodes connected by shared institutional investors to model information flow across these links. Fedyk (2024) developed a visibility-based diffusion model, treating news articles as nodes and linking them to investors according to Bloomberg terminal exposure, to map how information access shapes the effective network of information transmission. These approaches formalize information propagation in structured networks. The third stream are network finance frameworks. These provide theoretical foundations for understanding how investor linkages shape market dynamics. Ozsoylev and Walden (2011), using a noise trader equilibrium model, illustrated how information linkages among investors influence trading volume and the informational content of prices. Ozsoylev et al. (2014) extended the approach to dynamic networks, combining institutional ownership and transaction data to build evolving adjacency matrices. Walden (2019) developed a noisy rational expectations model of dynamic information networks, showing how information propagation through investor linkages shapes trading behavior, profits, and volatility. These studies formalize how network structures influence trading behavior.

Building upon the literature reviewed above, we identify several gaps that warrant further investigation. First, while prior studies have measured investor attention using aggregate proxies such as search volume, post frequency, and social media discussion intensity, relatively few have considered the relational structure of attention, that is, how interactions among investors form networks that influence the dissemination of attention. Second, existing research on information diffusion has largely focused on public news, institutional ownership links, or media visibility, and less on retail investor-driven interactions on social media, where decentralized discussion patterns generate complex information-flow networks. Third, while network finance frameworks have provided important theoretical foundations, less is empirically known about how information flow networks influence trading behavior and price discovery through channels such as opinion convergence and emotional resonance. To address these gaps, this study investigates the systemic impact of investor interaction-based information flow networks on market dynamics, combining perspectives from attention, diffusion, and network theory. By doing so, it aims to provide new theoretical and empirical insights into how social media interactions among investors shape volatility, trading, and price discovery in the market.

3. Research hypotheses

In social networks, the centrality of a node—such as a particular stock—within an information flow network typically indicates its greater capacity to receive and disseminate information, as well as to attract the attention of other nodes (Zhao et al., 2024). In a financial context, this implies that firms located at the core of the information diffusion network are more likely to trigger or gather market sentiment. Figure 2 presents the research framework discussed below.

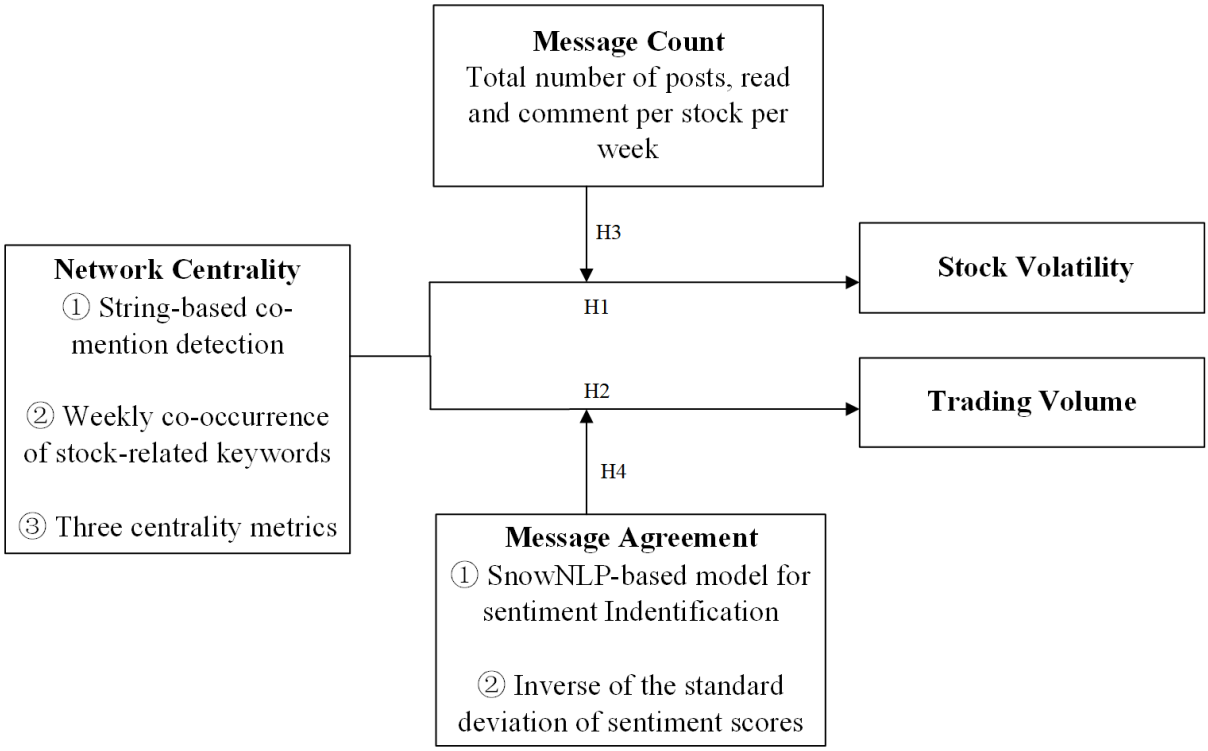


Figure 2. Research framework.

3.1. Hypothesis 1: Network centrality and stock price volatility

Prior research has offered multiple perspectives on the relationship between network structure and price volatility. Some studies (Zhai et al., 2024; Vamossy, 2024) showed that higher network density and centrality lead to faster information flow, which increases the speed and strength of price reactions. This is seen as a sign of improved information efficiency in the market. But other research questions this "volatility amplification" result. Huang et al. (2023) found that companies with high centrality on Xue Qiu had lower volatility. This suggests that not all transmitted information is valid—some may act as amplifiers of sentiment. Szczygielski et al. (2024) also pointed out that social media is full of naive and emotional content, which can cause information overload and bias, reducing the market's ability to respond correctly.

These studies show that network structure affects volatility, but the direction is unclear. The key is whether centrality spreads true information or just noise. In the context of this study, given the large number of retail investors and diverse information channels in the A-share market, we posit that companies with higher network centrality have greater information diffusion efficiency. As a result, their stock prices are more responsive to circulating information, leading to greater return volatility. Therefore, we propose Hypothesis 1:

H1. In the investor interaction network, the higher the network centrality of a listed company, the greater the stock price volatility, and this effect has a certain degree of persistence.

3.2. Hypothesis 2: Network centrality and trading volume

Unlike price volatility, trading volume is a better indicator of the persistence of investor behavioral responses. Prior studies suggest that firms located at the center of information networks are more visible and their information more easily disseminated, which triggers more frequent trading activity (Warkulat and Pelster, 2024; Messaoud et al., 2023). Especially on social media platforms, frequently discussed companies tend to attract more attention from retail investors and quickly generate a short-term "consensus for trading".

The drivers behind this increased trading activity can be understood through two key mechanisms. (1) Attention-driven mechanism: Investors are more likely to trade companies that are "salient" or prominently visible (Reichenbach et al., 2024; Ayala et al., 2024). (2) Opinion convergence mechanism: Frequent interactions and discussions reduce information asymmetry and strengthen belief alignment among investors (Bahari et al., 2024; Liang and Ye, 2024; Ortega-Rodríguez et al., 2024). Therefore, we expect that companies with higher centrality in the information flow network will not only attract more trading activity in the current period but also maintain elevated trading volumes over time due to sustained market attention. Thus, we propose Hypothesis 2:

H2. In the investor interaction network, the higher the network centrality of a listed company, the greater the trading volume, and this effect has a certain degree of persistence.

3.3 Hypotheses 3 and 4: Channel hypothesis

Beyond the direct effects of network centrality on volatility and trading volume, investor interaction networks may influence market outcomes through two distinct mechanisms (Qian et al., 2025; Wen, 2025). First, firms occupying central positions in the information flow network are more frequently discussed and more visible to investors. This may accelerate the diffusion of information and its incorporation into prices, thereby generating short-term price responses. Second, repeated and widespread discussions around central firms may facilitate belief alignment among investors. Such opinion convergence may strengthen coordinated trading behavior and sustain elevated trading activity over time. Based on this reasoning, we propose the following potential channel hypothesis:

H3. The information flow network drives short-term price responses through a "prompt information incorporation channel".

H4. The information flow network enhances the persistence of trading behavior through a "consensus formation channel".

4. Data and methodology

4.1. Data collection and sample

We chose comment data from the East Money Stock Forum (https://www.eastmoney.com/default.html) as the data source, as this forum has the largest number of users and the greatest influence among the mainstream Internet social platforms in China. It gathers abundant information flow from investors. The China Stock Market & Accounting Research (CSMAR) database was another data source, which was used to obtain firm-level attributes.

Using a crawler written in RStudio, we obtained both structured and unstructured data of 5,258,185 row comment data from June 1, 2023, to May 31, 2024, on the East Money Stock Forum. Figure 1 presents an example page of the interaction platform.

Our sample consisted of CSI 300 constituents. To reduce noise and improve data stability, daily observations were aggregated to the weekly level. Additionally, all continuous variables were winsorized at the 1st and 99th percentiles to mitigate the influence of extreme values on the regression results. Table 1 shows that the mean of idiosyncratic return volatility (IRV) is 0.022, with a standard deviation of 0.015. The average trading volume is 195 million shares, with substantial variation, indicating notable differences in trading activity across CSI 300 constituent stocks. In terms of network structure variables, all three types of centrality measures exhibit high dispersion. Notably, eigenvector centrality (EC) has the largest standard deviation (0.157), suggesting significant differences in firms' information influence on social media platforms. The distributions of other control variables appear reasonable and are consistent with the sample characteristics reported in prior literature (e.g., Vozlyublennaia, 2013; Benkraiem, 2022).

Table 1. Descriptive statistics.

Variable	Label	Description	Median	Mean	Std. dev.
Dependent	IRV	Idiosyncratic return volatility	0.019	0.022	0.015
	Volume	Trading Volume	98710611	195067505.644	255055100
Independent	DC	Degree centrality	0.017	0.025	0.029
	BC	Betweenness centrality	0.001	0.005	0.010
	EC	Eigenvector centrality	0.087	0.138	0.157
Control	Size	Firm size	77068939.330	129129019.955	175256600
	MB	Market-to-book ratio	1.986	2.868	2.661
	Return	Return	−0.004	−0.002	0.045
	IO	Institutional ownership	5.852	6.308	4.363
	Turnover	Excess turnover	0.006	0.008	0.009
	Volatility	Price volatility	0.157	0.184	0.092

4.2. Variable measurement

(1) Dependent variables. This paper looks at how investor networks affect two market behaviors—price volatility and trading activity—and defines dependent variables based on these. Following prior studies, this paper uses idiosyncratic return volatility (IRV) to measure stock price volatility and trading volume to reflect market activity (Ang et al., 2006; Xu et al., 2024). IRV captures the residual risk of an individual stock after accounting for systematic risk. A higher IRV indicates that the stock is more sensitive to firm-specific information shocks. Meanwhile, greater trading volume

typically reflects more frequent trading activity, indicating heightened investor attention and enthusiasm toward the stock.

(2) Building the social media information flow network. To identify "information interaction relationships" among stocks, this study follows the method used by Rossi et al. (2018) to measure network centrality among institutional investors and constructs 48 weekly investor interaction-based information flow networks using forum post content from the Eastmoney Stock Forum. Specifically, the more frequently stock j is mentioned in the discussion forum of stock i, the stronger the information interaction between the two stocks is considered to be. For example, on February 7, 2024, a user named "FriendqXUuHK" posted in the Ping An Bank forum: "The true short-selling forces are hiding behind Ping An Bank and ZTE." This post clearly indicates information flow between Ping An Bank (000651) and ZTE Corporation (000063). Another example comes from a January 3, 2024, post in the forum of Zhongji Xuchuang by the user "Ordinary Piggy": "AI concept stocks are falling sharply—Zhongji Xuchuang and iFlytek both dropped more than 5% in the afternoon." This indicates an interaction between Zhongji Xuchuang (300308) and iFlytek (002230). The specific text analysis approach involves string detection: If a post in stock i's forum mentions the name or code of stock j, an information link is identified between i and j. The edge weight between stocks i and j is defined as the number of such mentions, and this value is recorded as element Aijt in the network's adjacency matrix. For shortest-path-based measures, such as closeness centrality and betweenness centrality, the interaction strength is transformed into a distance measure by taking its reciprocal (Zhu, 2020). The underlying logic is that co-mentions of other stocks in the discussion of a focal stock reflect retail investors' cognitive associations formed through information search, attention allocation, and trading comparisons, thereby providing an observable basis for constructing investor-driven information-flow links (Ge et al., 2025; Huang et al., 2025; Schwenkler, 2025).

(3) Independent Variables. To comprehensively capture the structural characteristics of investor interaction-based information flow networks, we extract two types of core metrics: a) global network characteristics and b) local network centrality measures. These are used to describe how far, fast, and strongly information spreads.

a) Global network characteristics (global indicators). Following approaches in social network analysis and to ensure model robustness, we select three indicators to describe global network features: network density, network hierarchy, and network efficiency (Albery et al., 2024; Alinaghian et al., 2021). Network density is the number of real links divided by the number of all possible links. It shows how widely investors interact. A higher density indicates tighter network connectivity and more complete information flows. Network hierarchy shows if the network has a core–periphery structure. If few nodes are mutually reachable, the network has a clearer hierarchy and more peripheral nodes. Network efficiency measures how short and non-redundant the paths are. Shorter, cleaner paths indicate higher efficiency in information transmission. Let N denote the number of nodes in the network, K the number of symmetrically reachable node pairs, and M the number of redundant connections. The computation of the above metrics is defined as follows:

$$\mathrm{D} = \frac{\mathrm{L}}{\mathrm{N} \times (\mathrm{N} - 1)},$$

$$\mathrm{H} = 1 - \frac{\mathrm{K}}{\max(\mathrm{K})},$$

$$\mathrm{E} = 1 - \frac{\mathrm{M}}{\max(\mathrm{M})}.$$

In information flow networks dominated by retail investors, high values of density, hierarchy, and efficiency typically indicate close connections, a well-structured hierarchy, and efficient transmission paths.

b) Local network characteristics (centrality measures). To measure the structural centrality of individual listed firms within the social information network, we adopt three commonly used node-level centrality measures: degree centrality, closeness centrality, and betweenness centrality (Yu et al., 2024; Bi et al., 2021). Degree centrality counts how many direct links a node has—showing how often it is mentioned. More links mean the company is more visible in the network. Closeness centrality reflects the average shortest path length from one node to all others, indicating a firm's ability to access and disseminate information. Shorter paths mean more efficient information flow. Betweenness centrality measures the extent to which a node lies on the shortest paths between other nodes. It reflects the firm's role as a "bridge" in the network and its importance as a potential channel for information transmission. The indicators are defined as follows:

$$DC_j = C_1/(2N-2),$$

$$BC_j = 1000 \times \frac{\sum_{k\neq i;j\neq k;j\neq i} P_j(k,i)/P(k,i)}{(N-1)(N-2)/2},$$

$$EC_j = \lambda \sum_{i=1}^{N} A_{ij}\, E_i.$$

Let C_1 be the number of links of node j, N the total nodes, $P(k,i)$ the total shortest paths between k and i, $P_j(k,i)$ the number of those passing through j, and λ the largest eigenvalue of the adjacency matrix A_{ij}. In general, higher centrality means the company is more visible, spreads more information, and has greater potential impact on the market.

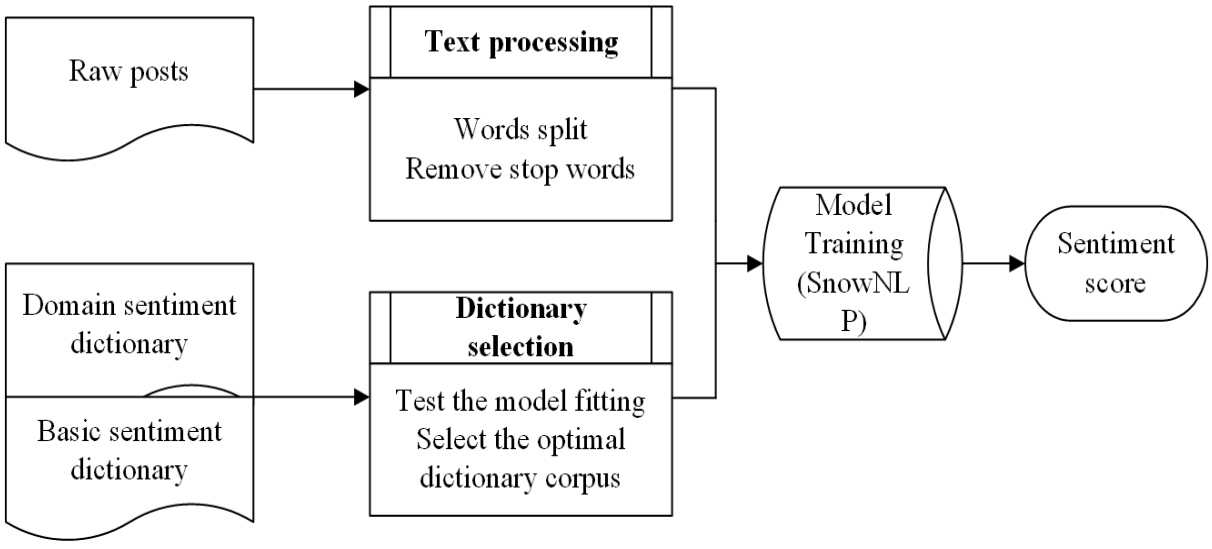


Figure 3. Opinion convergence measurement: text sentiment analysis process.

(4) Mediating variable. The intensity of information interaction is measured by the number of forum posts, reads, and comments, which serves as a proxy for market attention and the intensity of

information transmission. Opinion convergence is measured as the inverse of the standard deviation of sentiment scores derived from investor comments. Figure 3 shows the process of sentiment analysis, which embeds manual detection and automatic classification of the sentiment. To sum up, these sentiment scores are quantified using Python's SnowNLP model, which is trained on a composite financial sentiment lexicon, integrating resources such as SnowNLP's native corpus, HowNet, the NTU Simplified Chinese Sentiment Dictionary, BosonNLP, and CFSD.

4.3. Baseline comparison

This study adopts SnowNLP as the primary tool for sentiment quantification. SnowNLP is a lightweight, lexicon-based Chinese natural language processing toolkit that leverages word matching and naïve Bayes classifiers. The choice of SnowNLP is motivated by two considerations: (1) it offers high computational efficiency, which is particularly important for large-scale time-series sentiment extraction tasks; and (2) it provides transparent and interpretable results, making it more suitable for empirical research in financial markets where interpretability is critical.

To further assess the reliability of the sentiment classification, we conduct a manual validation exercise. Specifically, we randomly select 1000 comments from the Eastmoney Stock Forum and manually label them as positive or negative. The SnowNLP classification results are then compared with the manual labels, and standard classification metrics, including accuracy, precision, and F1-score, are reported.

Table 2. Validation results of sentiment classification models.

Model	Accuracy	Precision	F1-score
Built-in SnowNLP	0.699	0.763	0.717
Basic-corpus SnowNLP	0.705	0.775	0.726
Financial-domain SnowNLP	0.714	0.796	0.734

The results are reported in Table 2. Overall, the domain-specific SnowNLP model performs best among the three alternatives, with an accuracy of 0.714, precision of 0.796, and F1-score of 0.734. Compared with the built-in SnowNLP model and the basic-corpus model, the financial-domain-corpus model achieves modest improvements across all evaluation metrics, suggesting that incorporating financial-domain textual information helps improve sentiment classification performance.

Although SnowNLP is computationally efficient and interpretable for large-scale Chinese sentiment analysis, it may not fully capture slang, irony, implicit emotions, and domain-specific financial expressions in investor comments. Compared with more advanced NLP methods, such as domain-specific FinBERT, Chinese financial BERT models, or large language model–based text embeddings, SnowNLP has weaker contextual understanding. Therefore, the sentiment-based variables in this study are interpreted as noisy but usable proxies for investor sentiment.

4.4. Empirical model

Before estimating the baseline panel regression models, we first conduct Granger causality tests to examine the temporal direction between investor interaction networks and market outcomes. This step is motivated by the potential reverse causality concern that market volatility and trading activity may stimulate online discussions and increase network centrality. If such reverse causality is detected,

the baseline regressions further adopt a lagged structure, to mitigate potential endogeneity concerns (Bellemare, 2017; Leszczensky and Wolbring, 2022).

Based on this temporal-direction analysis, we then construct baseline panel regression models to further test the effects of investor interaction networks on market outcomes. Specifically, to test Hypothesis 1—whether the investor interaction-based information flow network affects stock price volatility—and to provide additional evidence for causal inference, this study specifies the following panel regression models:

Model 1:

$$\mathrm{IRV}_{it} = \beta_0 + \beta_1 \mathrm{Centrality}_{it} + \beta_2 \mathrm{Controls}_{it} + \mathrm{Week}_t + \mathrm{Firm}_i + \varepsilon_{it}$$

Model 2:

$$\mathrm{FIRV}_{it} = \beta_0 + \beta_1 \mathrm{Centrality}_{it} + \beta_2 \mathrm{Controls}_{it} + \mathrm{Week}_t + \mathrm{Firm}_i + \varepsilon_{it}$$

where IRV_{it} denotes idiosyncratic return volatility of firm i in week t, and FIRV_{it} denotes future idiosyncratic return volatility, measured as the cumulative idiosyncratic volatility over the following four weeks. The use of FIRV_{it} introduces a forward-looking lag structure, which helps examine whether current network centrality has predictive power for subsequent volatility and therefore provides additional support for causal inference. The key independent variables are local network features, including degree centrality, betweenness centrality, and eigenvector centrality. Subscripts i and t denote firm and week, respectively. Following prior studies (Hutton, 2005; Vozlyublennaia, 2013; Wang, 2018; Benkraiem et al., 2022), the model controls for industry fixed effects and includes the following firm-level control variables: firm size (Size), market-to-book ratio (MB), stock return (Return), institutional ownership (IO), and excess turnover (Turnover). ε_{it} is the random error term.

To test Hypothesis 2—whether the investor interaction-based information flow network influences trading volume—and to provide additional evidence for causal inference, this study establishes the following panel regression models:

Model 3:

$$\mathrm{Volume}_{it} = \beta_0 + \beta_1 \mathrm{Centrality}_{it} + \beta_2 \mathrm{Controls}_{it} + \mathrm{Week}_t + \mathrm{Firm}_i + \varepsilon_{it}$$

Model 4:

$$\mathrm{Future_volume}_{it} = \beta_0 + \beta_1 \mathrm{Centrality}_{it} + \beta_2 \mathrm{Controls}_{it} + \mathrm{Week}_t + \mathrm{Firm}_i + \varepsilon_{it}$$

where Volume_{it} denotes trading volume of firm i in week t, and $\mathrm{Future_volume}_{it}$ denotes future trading volume, measured as the cumulative trading volume over the following four weeks. The use of $\mathrm{Future_volume}_{it}$ introduces a forward-looking lag structure, which helps examine whether current network centrality has predictive power for subsequent volume and therefore provides additional support for causal inference. The key independent variables again include degree centrality, betweenness centrality, and eigenvector centrality. Subscripts i and t denote firm and week, respectively. In addition to controlling for industry fixed effects, we include the following control variables: firm size (Size), market-to-book ratio (MB), stock return (Return), institutional ownership (IO), and price volatility (IRV). ε_{it} is the random error term.

5. Empirical analysis

5.1. Analysis of information flow network characteristics

To gain a comprehensive understanding of the structural characteristics and evolutionary patterns of investor interaction-based information flow networks, we conducted both visualization and time-series analyses on the global and local indicators across 48 weekly network periods, as shown in Figure 4.

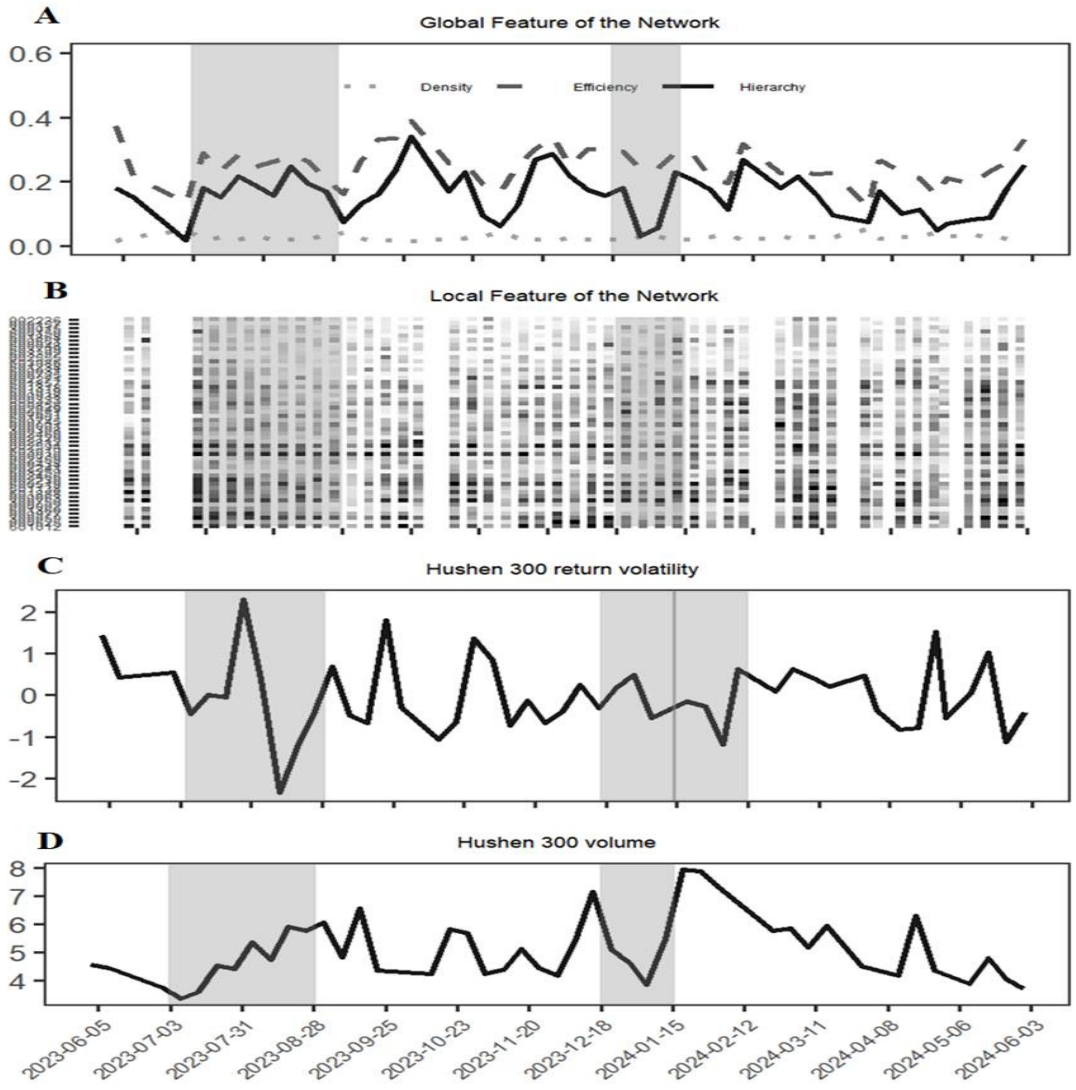


Figure 4. Global and local features of the information flow network and time series of price fluctuations and trading volume.

Figure 4A shows three global metrics: density, hierarchy, and efficiency. The results show that the average network density is 0.02, indicating weak overall connectivity and limited information diffusion among retail investors. The hierarchy index is 0.16, suggesting a relatively flat network structure with no obvious information "pyramid". The network efficiency score is 0.25, implying that, despite sparse connectivity, redundant paths provide a degree of structural stability. These findings suggest that while

the frequency of investor information exchange is limited, the network is structurally stable and unlikely to experience systemic collapse-type fluctuations.

Figure 4B visualizes the eigenvector centrality of the top 50 stocks by average comment volume, revealing substantial variation in node prominence. Some stocks, such as Zhongji Xuchuang and iFlytek, exhibit strong centralization, indicating their greater topic traction and influence in information dissemination on social media. Further comparison using a heatmap reveals a strong cross-sectional alignment between comment volume and node centrality. Stocks with more comments tend to exhibit higher centrality in the information flow network. This finding confirms a positive relationship between the frequency of investor discussions and a firm's central position in the information network, reinforcing the validity of centrality as a representative measure of information flow and attention.

Figures 4A–4D illustrate the time-series evolution of network metrics, return volatility, and trading volume. Several periods exhibit clear synchronous or lagged response patterns: From July 3 to August 28, 2023, both network density and efficiency first increased and then declined, accompanied by similar fluctuations in return volatility and trading volume. From December 18, 2023, to January 15, 2024, network hierarchy and efficiency reached a trough, during which trading volume dropped simultaneously, followed by a significant increase in volatility after a time lag. These patterns suggest that the structure of information diffusion on social media not only correlates with market behavior contemporaneously but may also serve as a leading indicator in certain periods, implying predictive potential in network features.

5.2. Granger causality analysis

Bidirectional Granger causality tests were conducted to examine the temporal predictive relationship between network centrality and market outcomes. Table 3 reports the Granger causality results between three centrality measures and two market variables, namely idiosyncratic return volatility (IRV) and trading volume (Volume), under different lag specifications.

Table 3. Granger causality tests between network centrality and market outcomes.

Hypothesis	Lag 1	Lag 2	Lag 3	Lag 4
IRV ⇏ DC	6.191***	5.816***	5.193***	4.508***
DC ⇏ IRV	1.439	3.189***	3.998***	2.521**
IRV ⇏ BC	4.882***	5.999***	4.074***	5.234***
BC ⇏ IRV	2.905***	0.422	1.933*	3.205***
IRV ⇏ EC	4.080***	4.367***	3.768***	3.440***
EC ⇏ IRV	1.751*	0.716	1.035	1.157
Volume ⇏ DC	11.725***	8.681***	7.995***	6.998***
DC ⇏ Volume	4.187***	3.407***	3.240***	2.768***
Volume ⇏ BC	3.978***	6.056***	5.781***	7.009***
BC ⇏ Volume	3.536***	2.725***	5.541***	4.704***
Volume ⇏ EC	9.114***	6.549***	6.209***	4.895***
EC ⇏ Volume	2.281**	1.732*	3.287***	2.795***

Note: This table reports Ztilde statistics from the Dumitrescu-Hurlin panel Granger non-causality tests. The null hypothesis is that the variable on the left-hand side of "⇏" does not Granger-cause the variable on the right-hand side. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

The findings in Table 3 indicate a bidirectional predictive relationship between network centrality and market outcomes. On the one hand, several network centrality measures significantly Granger-cause idiosyncratic return volatility and trading volume. For example, degree centrality significantly predicts IRV at lags 2–4 and predicts trading volume across all lag specifications. Betweenness centrality also shows significant predictive power for both IRV and trading volume in most lag specifications, while eigenvector centrality exhibits more consistent predictive effects on trading volume than on IRV. On the other hand, the reverse direction is also statistically significant in many cases. This finding indicates that market fluctuations and trading activity may also stimulate online discussions and reshape firms' positions in the investor interaction network. Therefore, the relationship between social media information-flow networks and financial market outcomes should be understood as a dynamic feedback process rather than a strictly one-way causal channel.

Accordingly, the baseline panel regressions adopt a lagged structure to mitigate potential endogeneity concerns. Specifically, current network centrality is used to explain future volatility and future trading volume, which helps reduce the influence of contemporaneous reverse causality from market movements to online discussions.

5.3. Regression analysis

Fixed effects models were employed to test Hypotheses 1 and 2. Tables II and III report the effects of network centrality on idiosyncratic return volatility (IRV) and trading volume (Volume), respectively.

Table 4. Empirical results of network centrality and price fluctuation.

	IRV			Future IRV		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.013**			0.001		
	(0.006)			(0.025)		
BC		0.015**			−0.013	
		(0.006)			(0.024)	
EC			−0.007			−0.084***
			(0.006)			(0.023)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.029	0.027	0.029	0.055	0.055	0.056

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

As shown in Table 4, degree centrality (DC) and betweenness centrality (BC) both exhibit significantly positive contemporaneous effects on IRV, with coefficients of 0.013 and 0.015, respectively (both significant at the 5% level). These findings provide empirical support for Hypothesis 1, indicating that firms occupying structurally central positions in the information flow network are more sensitive to shifts in investor-generated information. In addition, eigenvector centrality (EC) shows a negative association with future volatility (FIRV), with a coefficient of −0.084 significant at

the 1% level. Economically, a one-unit increase in DC and BC is associated with 0.013 and 0.015 higher contemporaneous IRV, respectively, while a one-unit increase in EC is associated with 0.084 lower future IRV. This result suggests that although highly central firms tend to experience greater short-term price volatility, the market tends to absorb their information more efficiently, leading to a reduction in volatility over time—a dynamic that reflects the rapid digestion of information premiums.

Table 5. Empirical results of network centrality and trading volume.

	Volume			Future volume		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.250***			0.345***		
	(0.005)			(0.018)		
BC		0.171***			0.242***	
		(0.005)			(0.018)	
EC			0.193***			0.304***
			(0.005)			(0.017)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.029	0.027	0.029	0.055	0.055	0.056

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

The results presented in Table 5 indicate that all three network centrality measures exhibit highly significant positive effects on contemporaneous trading volume. For instance, the coefficient for degree centrality (DC) is 0.250, with all measures significant at the 1% level. These findings provide strong support for Hypothesis 2, implying that firms occupying more central positions in the information flow network tend to attract greater investor attention and trading activity due to broader information dissemination. This underscores a robust link between social influence and market participation. Moreover, the coefficients of all three centrality measures on future trading volume are also significantly positive, indicating persistent effects over time. In terms of economic magnitudes, a one-unit increase in degree, betweenness, and eigenvector centrality is associated with 0.250, 0.171, and 0.193 higher contemporaneous trading volume, and 0.345, 0.242, and 0.304 higher future trading volume, respectively. This suggests that firms with higher network centrality not only trigger stronger trading responses but also maintain elevated investor engagement, reflecting sustained trading behavior in response to their social media presence.

It should be noted that the adjusted R2 values in the baseline regressions are relatively low. This is common in short-horizon financial-market regressions, especially when the dependent variables are relatively high-frequency market outcomes such as volatility, returns, or trading activity. For example, recent empirical studies on short-horizon price pressure and investor behavior also reported low explanatory power in predictive regressions, with adjusted R2 values around 0.01–0.03 in some specifications (Harvey et al., 2025; Jiang et al., 2026). More broadly, recent studies showed that short-term market outcomes are jointly affected by attention spillovers, disagreement, news diffusion, liquidity shocks, and social-media signals (Chen et al., 2023; Cookson et al., 2024). Therefore, the

limited explanatory power of a parsimonious panel model is understandable in studies of relatively high-frequency financial market outcomes.

5.4. Further analysis

To provide suggestive evidence on the potential channels through which investor interaction networks are associated with market outcomes, we constructed two channel variables: message count and future message count, which capture the intensity of information interaction among investors in the current and subsequent periods, respectively; and message agreement and future message agreement, which reflect the degree of consensus in investor opinions in the current and future periods.

Table 6. Empirical results of network centrality and stock forum comments and opinion convergence degree.

	Panel A: Comment count					
	Message count			Future message count		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.506***			−0.685***		
	(0.013)			(0.015)		
BC		0.178***			−0.430***	
		(0.013)			(0.015)	
EC			0.265***			−0.411***
			(0.012)			(0.014)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.029	0.027	0.029	0.055	0.055	0.056
	Panel B: Message agreement					
	Message agreement			Future message agreement		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.129***			0.351***		
	(0.012)			(0.026)		
BC		0.150***			0.208***	
		(0.012)			(0.025)	
EC			0.103***			0.230***
			(0.011)			(0.024)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.029	0.027	0.029	0.055	0.055	0.056

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

Regression results are reported in Table 6, Panel A. Columns (1)–(3) show that all three centrality measures—degree centrality (DC), betweenness centrality (BC), and eigenvector centrality (EC)—are positively and significantly associated with contemporaneous comment volume, indicating that firms with higher centrality attract more discussion in the current period. In contrast, Columns (4)–(6) demonstrate that centrality has a significantly negative effect on future comment volume, suggesting that investor attention toward these firms fades rapidly over time. These findings support the "prompt information incorporation channel" (Hypothesis 3): Firms occupying core positions in the information flow network benefit from shorter and faster information transmission paths, leading to quicker price

incorporation and elevated short-term volatility. However, as the relevant information is rapidly absorbed by the market, its impact diminishes quickly, making it difficult for volatility to persist.

Table 6, Panel B presents the regression results. Columns (1)–(3) show that all three network centrality measures are significantly positively associated with contemporaneous opinion convergence, indicating that firms with higher centrality facilitate more aligned investor expectations during the current period. Furthermore, Columns (4)–(6) reveal that this association extends into future periods, suggesting that expectation alignment tends to persist over time. These findings support the "consensus formation channel" (Hypothesis 4): Companies in central network positions are more likely to generate consensus among investors, which in turn sustains coordinated trading behavior over an extended horizon.

6. Robustness analysis

6.1. Robustness analysis based on alternative network construction

We further constructed an alternative information-flow network based on sentiment correlations. Specifically, we first used SnowNLP to calculate the daily sentiment score of investor comments for each stock. Then, for each week, we computed the pairwise sentiment correlations over a rolling four-week window. Positive correlation coefficients were retained as edge weights to construct a weighted sentiment-correlation network, so that stronger correlations indicate stronger sentiment linkages between stocks. Based on this alternative network, we recalculated the centrality measures and re-estimated the baseline regression models.

Table 7. Robustness analysis using alternative network construction: effects on price volatility.

	IRV			Future IRV		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	−0.006			−0.058		
	(0.009)			(0.038)		
BC		0.001			−0.008	
		(0.005)			(0.023)	
EC			0.019***			−0.047***
			(0.006)			(0.024)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.050	0.050	0.051	0.058	0.061	0.057

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

As shown in Table 7, the effects of network centrality on price volatility are partly consistent with the baseline results. Specifically, EC is significantly positively associated with contemporaneous IRV and significantly negatively associated with future IRV, indicating that stocks occupying more influential positions in the sentiment-correlation network tend to experience stronger short-term volatility but lower subsequent volatility. The coefficients of DC and BC are not statistically significant, suggesting that the volatility effect varies across different centrality measures.

Table 8. Robustness analysis using alternative network construction: effects on trading volume.

	Volume			Future volume		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.056***			0.134***		
	(0.008)			(0.028)		
BC		0.037***			0.114***	
		(0.005)			(0.017)	
EC			0.043***			-0.010
			(0.005)			(0.017)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.191	0.192	0.192	0.025	0.027	0.024

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

Table 8 reports the results for trading volume. The coefficients of DC, BC, and EC are all positive and highly significant for contemporaneous trading volume, indicating that stocks with higher centrality in the sentiment-correlation network attract stronger current trading activity. For future trading volume, DC and BC remain significantly positive, while EC is not statistically significant. These findings suggest that the main conclusion regarding trading activity remains robust under the alternative network construction.

6.2. Robustness analysis based on alternative independent variables

To examine whether the baseline findings are sensitive to the measurement of market outcomes, we further conducted robustness tests using alternative dependent variables. Specifically, we replaced idiosyncratic return volatility with the absolute value of stock returns, |Return|, as an alternative measure of price fluctuation. In addition, we replaced trading volume with turnover as an alternative measure of trading activity. The baseline regression models were then re-estimated using these alternative measures and their future counterparts.

As shown in Table 9, when |Return| is used as the alternative volatility measure, the coefficients of DC, BC, and EC are all positive and highly significant for contemporaneous price fluctuation. For future |Return|, DC and EC are significantly negative, while BC is not statistically significant. These results are generally consistent with the baseline findings, suggesting that network centrality is associated with stronger contemporaneous price responses, while its effect on future volatility is weaker and varies across centrality measures.

Table 9. Robustness analysis using alternative volatility measure.

	|Return|			Future |Return|		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.092***			-0.036*		
	(0.008)			(0.018)		
BC		0.062***			0.010	
		(0.007)			(0.016)	
EC			0.067***			−0.064***
			(0.007)			(0.016)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.074	0.071	0.072	0.020	0.021	0.021

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

Table 10 presents the results using turnover as the alternative measure of trading activity. The coefficients of DC and BC are significantly positive for both contemporaneous and future turnover. EC is not significant for contemporaneous turnover but becomes significantly positive for future turnover. These findings indicate that firms with higher network centrality tend to experience stronger subsequent trading activity. Overall, the robustness tests based on alternative dependent variables provide additional support for the baseline conclusions and suggest that the main results are not driven by a specific measurement of volatility or trading activity.

Table 10. Robustness analysis using alternative trading volume measure.

	Turnover			Future turnover		
	(1)	(2)	(3)	(4)	(5)	(6)
DC	0.076***			0.295***		
	(0.013)			(0.024)		
BC		0.052***			0.212***	
		(0.008)			(0.022)	
EC			0.013			0.249***
			(0.008)			(0.021)
Controls	yes	yes	yes	yes	yes	yes
Firm and week FE	yes	yes	yes	yes	yes	yes
N	14400	14400	14400	14400	14400	14400
Adj.R^2	0.015	0.014	0.002	0.002	0.007	0.004

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

6.3. Lagged regression analysis

To further mitigate potential endogeneity concerns and examine the temporal robustness of the baseline results, additional lagged regression analyses were conducted from one-week to four-week horizons. Specifically, current network centrality was used to predict future market outcomes at t+1, t+2, t+3, and t+4.

As shown in Table 11, the effects of network centrality on future IRV are mainly reflected in degree centrality and eigenvector centrality. Specifically, DC and EC are significantly negatively associated with future IRV across most lag specifications, while the coefficient of BC is not statistically significant. These results suggest that the impact of investor interaction networks on stock price volatility is relatively short-term and varies across different centrality measures.

Table 11. Robustness analysis: Lagged regression results for future IRV.

Independent variables	IRV_{t+1}	IRV_{t+2}	IRV_{t+3}	IRV_{t+4}
Dependent Variables				
DC	−0.012**	−0.023***	−0.029***	−0.032***
	(0.006)	(0.006)	(0.006)	(0.006)
BC	0.002	−0.003	−0.006	−0.007
	(0.006)	(0.006)	(0.006)	(0.006)
EC	−0.013**	−0.019***	−0.021***	−0.035***
	(0.006)	(0.006)	(0.006)	(0.006)
Controls	Yes	Yes	Yes	Yes
Firm and week FE	Yes	Yes	Yes	Yes
N	14400	14400	14400	14400
Adj. R^2	0.043	0.065	0.069	0.076

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

Table 12 presents the lagged regression results for future trading volume. The coefficients of DC, BC, and EC remain positive and highly significant from t+1 to t+4, indicating that firms with higher network centrality continue to attract stronger trading activity in subsequent periods. Compared with the results for IRV, the effect on trading volume is more stable and persistent. Overall, the lagged regression results confirm that network centrality has predictive power for future market outcomes and provide additional support for the baseline conclusions.

Table 12. Robustness analysis: Lagged regression results for future trading volume.

Independent variables Dependent Variables	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
DC	0.147***	0.092***	0.082***	0.056***
	(0.006)	(0.006)	(0.006)	(0.006)
BC	0.112***	0.067***	0.051***	0.040***
	(0.005)	(0.006)	(0.006)	(0.006)
EC	0.116***	0.071***	0.070***	0.051***
	(0.005)	(0.021)	(0.005)	(0.006)
Controls	Yes	Yes	Yes	Yes
Firm and week FE	Yes	Yes	Yes	Yes
N	14400	14400	14400	14400
Adj. R^2	0.031	0.024	0.017	0.011

Note: *, **, and *** denote that the estimated values of the parameters are statistically significant at the 10%, 5%, and 1% levels, respectively; the values in parentheses are robust standard errors.

7. Conclusions

7.1. Major findings

In this study, several notable findings are obtained. Compared to prior literature, we find that the structure of our information flow network is relatively sparse but stable and exhibits certain predictive capabilities. The investor interaction network shows low overall density and an inconspicuous hierarchical structure yet contains redundant links that contribute to structural robustness. Time-series analysis reveals that network structural indicators are both contemporaneously and lag-correlated with return volatility and trading volume, indicating that such networks can partially predict market trends.

Innovatively, the empirical findings in our work reveal that network centrality exerts a "dual effect" on market behavior. Firms positioned at the core of the information network are more likely to attract attention and facilitate information diffusion, resulting in faster price responses and short-term increases in volatility, though such effects tend to fade quickly. At the same time, broad-based discussions around central firms foster expectation alignment among investors. This opinion convergence enhances trading activity and sustains higher trading volume over longer periods, underscoring the role of central firms in both information diffusion and belief coordination.

Further, our findings emphasize that the speed of information absorption and the degree of opinion convergence together form a unified mechanism. Mechanism tests show that firms with higher network centrality attract more intense discussion, leading to faster market digestion of information and stronger short-term price reactions. Simultaneously, these firms promote greater consensus among investors, which drives persistent trading activity. This explains the observed divergence between the volatility and volume responses, offering new insights into how information flow networks shape market dynamics through differentiated yet complementary channels.

7.2. Theoretical implications

From a theoretical perspective, this study contributes to the literature by empirically establishing a link between investor interaction-based information flow on social media and financial market volatility. While prior research has often focused on the effect of investor sentiment, our study constructs investor networks and analyzes their temporal dynamics. By introducing network centrality and opinion convergence as mediators, we provide a novel analytical lens to understand how decentralized social communication influences aggregate market outcomes (Mikhaylov et al, 2024; Mikhaylov, 2021).

Second, this study extends the literature on information diffusion and market microstructure by highlighting the role of interaction intensity and sentiment alignment. Most existing works treat social media as a unidirectional sentiment indicator, whereas our findings show that investor-driven communication patterns—including network topology and convergence of beliefs—play an essential role in shaping market behavior. This distinction not only enriches theoretical perspectives on social information processing but also opens new pathways for integrating network science into behavioral finance.

7.3. Practical implications

Based on the findings above, this study proposes policy recommendations from the perspectives of listed companies, regulatory agencies, and social media platforms. For listed companies, engaging with investors on social media may help improve information communication and reduce information uncertainty. Firms could consider using social platforms to provide timely updates, clarify potential misunderstandings, and monitor changes in investor sentiment. Such efforts may support more informed trading and help reduce short-term volatility.

For regulatory agencies, information flow networks on social media could provide supplementary signals for understanding market sentiment and investor attention. Authorities might explore the use of network indicators—such as density, centrality, and hierarchy—to identify unusual patterns of investor attention. Early identification of anomalous activity could potentially assist in mitigating systemic risk.

For social media platforms, improving information delivery mechanisms and promoting high-quality discussions may contribute to a healthier online information environment for investors. Platforms may consider enhancing content classification, reducing the spread of misleading information, and exploring privacy-preserving cooperation with regulators where appropriate. These measures may be useful for improving the quality of online financial discussions.

7.4. Limitations and future research directions

Although this study provides empirical evidence on the role of information flow networks in financial markets, several limitations remain and open avenues for future research. First, a single-platform data source. This study relies primarily on data from the Eastmoney Stock Forum. Future research could incorporate data from multiple social media platforms such as Weibo, Xueqiu, and Zhihu to enhance external validity and generalizability (Li et al., 2026; Chen et al., 2026). Second, a limited time span. Due to data availability, the study examines network dynamics over a one-year

period. Extending the sample to cover longer time horizons and applying event-based analytical methods could provide deeper insights into the temporal effects of network structure on market responses (Chen et al., 2025). Third, a lack of heterogeneity analysis in information content. This study does not differentiate between types of information (positive, or negative). Future research could use topic modeling techniques to explore how different kinds of information diffuse through the network and generate heterogeneous impacts on asset pricing dynamics and trading intensity (Shaikh et al., 2022; Irshad et al., 2024). Fourth, market-specific external validity: This study focuses on the Chinese A-share market, where retail investors are highly active. Thus, the documented network effects may be more pronounced in retail-investor-driven markets than in markets dominated by institutional investors. Future research could extend the proposed framework to markets with different investor structures, such as the U.S. or European markets, to further assess the generalizability of the findings.

Author contributions

Zhongxiu Chen: conceptualization, methodology, software, data curation, formal analysis, visualization, writing—original draft preparation. Zhenghui Li: conceptualization, methodology, validation, supervision, project administration, writing—review and editing. Both authors have read and agreed to the published version of the manuscript.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by Natural Science Foundation of Guangdong Province (2024A1515012587).

Conflict of interest

Zhenghui Li is an editor-in-chief for Data Science in Finance and Economics and was not involved in the editorial review or the decision to publish this article. All authors declare no conflicts of interest in this paper.

References

Albery GF, Bansal S, Silk MJ (2024) Comparative approaches in social network ecology. *Ecol Lett* 27: e14345. https://doi.org/10.1111/ele.14345

Alinaghian L, Qiu J, Razmdoost K (2021) The role of network structural properties in supply chain sustainability: A systematic literature review and agenda for future research. *Supply Chain Manag* 26: 192–211. https://doi.org/10.1108/SCM-11-2019-0407

An J, Mikhaylov A, Chang T (2024) Relationship between the popularity of a platform and the price of NFT assets. *Financ Res Lett* 61: 105057. https://doi.org/10.1016/j.frl.2024.105057

Ang A, Hodrick RJ, Xing Y, et al. (2006) The cross-section of volatility and expected returns. *J Financ* 61: 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Ayala MJ, Gonzálvez-Gallego N, Arteaga-Sánchez R (2024) Google search volume index and investor attention in stock market: A systematic review. *Financ Innov* 10: 70. https://doi.org/10.1186/s40854-023-00606-y

Bahari NN, Bahaludin H, Ismail M, et al. (2024) Network, correlation, and community structure of the financial sector of Bursa Malaysia before, during, and after COVID-19. *Data Sci Financ Econ* 4: 362–387. https://doi.org/10.3934/DSFE.2024016

Bellemare MF, Masaki T, Pepinsky TB (2017) Lagged explanatory variables and the estimation of causal effect. *J Poli* 79: 949–963. https://doi.org/10.1086/690946

Benkraiem R, Goutte S, Saadi S, et al. (2022) Investor heterogeneity and negative skewness in stock returns: Evidence from institutional investors. *J Int Financ Mark Institutions and Money* 81: 101690. https://doi.org/10.1016/j.intfin.2022.101690

Bi J, Jin J, Qu C, et al. (2021) Temporal gravity model for important node identification in temporal networks. *Chaos Soliton Fract* 147: 110934. https://doi.org/10.1016/j.chaos.2021.110934

Chen X, An L, Wang Z, et al. (2023) Attention spillover in asset pricing. *J Financ* 78: 3515–3559. https://doi.org/10.1111/jofi.13281

Chen Z, Tian H, Li Z (2026) Online monitoring of financial market information-flow networks under external shocks: A rolling directed-ERGM and control-chart framework. *Mathematics* 14: 961. https://doi.org/10.3390/math14060961

Chen Z, Yin J (2025) Detecting jumps in stochastic volatility jump-diffusion models via the power variation approach. *AIMS Mathematics* 10: 19189–19216. https://doi.org/10.3934/math.2025858

Cookson JA, Lu R, Mullins W, et al. (2024) The social signal. *J Financ Econ* 158: 103870. https://doi.org/10.1016/j.jfineco.2024.103870

Du Q, Kwaramba CS, Li C, et al. (2023) A contest between users and marketers? The economic value of social media content for adverse events. *Inform Process Manag* 60: 103496. https://doi.org/10.1016/j.ipm.2023.103496

Fedyk A (2024) Front-page news: The effect of news positioning on financial markets. *J Financ* 79: 5–33. https://doi.org/10.1111/jofi.13287

Ge S, Li S, Zheng H (2025) Diamond cuts diamond: News co-mention momentum spillover prevails in China. *J Bank Financ* 171: 107356. https://doi.org/10.1016/j.jbankfin.2024.107356

Han B, Yang L (2013) Social networks, information acquisition, and asset prices. *Manage Sci* 59: 1444–1457. https://doi.org/10.1287/mnsc.1120.1678

Harvey CR, Mazzoleni MG, Melone A (2025) The unintended consequences of rebalancing. *NBER Working Paper* No. 33554. https://doi.org/10.3386/w33554

Hirshleifer D, Peng L, Wang Q (2025) News diffusion in social networks and stock market reactions. *Rev Financ Stud* 38: 883–937. https://doi.org/10.1093/rfs/hhae025

Huang C, Cao Y, Lu M, et al. (2023) Messages in online stock forums and stock price synchronicity: Evidence from China. *Account Financ* 63: 3011–3041. https://doi.org/10.1111/acfi.13005

Huang Y, Li Y, Li P, et al. (2025) The unseen guardian: News co-mention and stock market manipulation. *Financ Res Lett* 86: 108648. https://doi.org/10.1016/j.frl.2025.108648

Hutton AP (2005) Determinants of managerial earnings guidance prior to regulation fair disclosure and bias in analysts' earnings forecasts. *Contemp Account Res* 22: 867–914. https://doi.org/10.1506/6QUR-CR5M-AQQX-KX1A

Irshad AS, Zakir MN, Rashad SS, et al. (2024) Comparative analyses and optimizations of hybrid biomass and solar energy systems based upon a variety of biomass technologies. *Energ Convers Man-X* 23: 100640. https://doi.org/10.1016/j.ecmx.2024.100640

Jiang Z, Liu H, Peng C, et al. (2025) Investor memory and biased beliefs: Evidence from the field. *Q J Econ* 140: 2749–2804. https://doi.org/10.1093/qje/qjaf035

Leszczensky L, Wolbring T (2022) How to deal with reverse causality using panel data? Recommendations for researchers based on a simulation study. *Sociol Method Res* 51: 837–865. https://doi.org/10.1177/0049124119882473

Li Z, Chen Z, Huang Z (2026) Modelling the data-generating mechanism of China's commodity market by identifying hidden information flow regimes. *Financ Innov* 12: 28. https://doi.org/10.1186/s40854-025-00804-w

Liang Z, Ye Q (2024) Despite absolute information advantages, all investors incur welfare loss. *arXiv* preprint arXiv:2405.08822. https://doi.org/10.48550/arXiv.2405.08822

Liao G, Li Y, Wang M (2024) Contagion network of idiosyncratic volatility: Does corporate environmental responsibility matter? *Energy Econ* 129: 107168. https://doi.org/10.1016/j.eneco.2023.107168

Luo R, Zhao S, Zhou J (2023) Information network, public disclosure and asset prices. *Pac-Basin Financ J* 77: 101882. https://doi.org/10.1016/j.pacfin.2022.101882

Messaoud D, Ben Amar A, Boujelbene Y (2023) Investor sentiment and liquidity in emerging stock markets. *J Econom Adm Sci* 39: 867–891. https://doi.org/10.1108/JEAS-11-2020-0198

Mikhaylov A (2021) Development of Friedrich von Hayek's theory of private money and economic implications for digital currencies. *Terra Econ* 19: 53–62. https://doi.org/10.18522/2073-6606-2021-19-1-53-62

Mikhaylov A, Bhatti IM, Dinçer H, et al. (2024) Integrated decision recommendation system using iteration-enhanced collaborative filtering, golden cut bipolar for analyzing the risk-based oil market spillovers. *Comput Econ* 63: 305–338. https://doi.org/10.1007/s10614-022-10341-8

Mikhaylov A, Dinçer H, Yüksel S, et al. (2023) Bitcoin mempool growth and trading volumes: Integrated approach based on QROF Multi-SWARA and aggregation operators. *J Innov Knowl* 8: 100378. https://doi.org/10.1016/j.jik.2023.100378

Nguyen MT, Meesad P, Nguyen HS (2024) On a stock prediction aligned to natural language sentiments. In: *Proceedings of the 2024 8th International Conference on Natural Language Processing and Information Retrieval*, 395–400. https://doi.org/10.1145/3711542.3711597

Ortega-Rodríguez C, Vena-Oya J, Barreal J, et al. (2024) How to finance sustainable tourism: Factors influencing the attitude and willingness to pay green taxes among university students. *Green Financ* 6: 649–665. https://doi.org/10.3934/GF.2024025

Ozsoylev HN, Walden J (2011) Asset pricing in large information networks. *J Econ Theor* 146: 2252–2280. https://doi.org/10.1016/j.jet.2011.10.003

Ozsoylev HN, Walden J, Yavuz MD, et al. (2014) Investor networks in the stock market. *Rev Financ Stud* 27: 1323–1366. https://doi.org/10.1093/rfs/hht065

Qian H, Zhou H, Zhao Q, et al. (2024) MDGNN: Multi-relational dynamic graph neural network for comprehensive and dynamic stock investment prediction. *Proceedings of the AAAI Conference on Artificial Intelligence* 38: 14642–14650. https://doi.org/10.1609/aaai.v38i13.29381

Reichenbach F, Walther M (2025) Attention allocation of investors on social media: The role of prospect theory. *J Behav Financ* 26: 317–334. https://doi.org/10.1080/15427560.2024.2309145

Rossi AG, Blake D, Timmermann A, et al. (2018) Network centrality and delegated investment performance. *J Financ Econ* 128: 183–206. https://doi.org/10.1016/j.jfineco.2018.02.003

Schwenkler G, Zheng H (2025) News-driven peer co-movement in crypto markets. *J Corp Financ* 93: 102772. https://doi.org/10.1016/j.jcorpfin.2025.102772

Shaikh ZA, Kraikin A, Mikhaylov A, et al. (2022) Forecasting stock prices of companies producing solar panels using machine learning methods. *Complexity* 2022: 9186265. https://doi.org/10.1155/2022/9186265

Shi Y, Wang Y, Qu Y, et al. (2024) Integrated GCN-LSTM stock prices movement prediction based on knowledge-incorporated graphs construction. *Int J Mach Learn Cybern* 15: 161–176. https://doi.org/10.1007/s13042-023-01817-6

Szczygielski JJ, Charteris A, Bwanya PR, et al. (2024) Google search trends and stock markets: Sentiment, attention or uncertainty? *Int Rev Financ Anal* 91: 102549. https://doi.org/10.1016/j.irfa.2023.102549

Vamossy DF (2026) Social media emotions and market behavior. *Int Rev Financ Anal* 114: 105162. https://doi.org/10.1016/j.irfa.2026.105162

Vozlyublennaia N (2013) Do firm characteristics matter for the dynamics of idiosyncratic risk? *J Int Financ Mark Inst Money* 27: 35–46. https://doi.org/10.1016/j.intfin.2013.07.006

Walden J (2019) Trading, profits, and volatility in a dynamic information network model. *Rev Econ Stud* 86: 2248–2283. https://doi.org/10.1093/restud/rdy058

Wang J, Liao L, Zhong K, et al. (2025) MRRFGNN: Multi-relation reconstruction and fusion graph neural network for stock crash prediction. *Inform Sci* 689: 121507. https://doi.org/10.1016/j.ins.2024.121507

Wang LH (2018) Which firm characteristics affect foreign analyst coverage? Evidence from the Taiwan stock market. *J Math Financ* 8: 64–85. https://doi.org/10.4236/jmf.2018.81006

Wang MC, Chang T, Mikhaylov A, et al. (2024) A measure of quantile-on-quantile connectedness for the US treasury yield curve spread, the US Dollar, and gold price. *North Am J Econ Financ* 74: 102232. https://doi.org/10.1016/j.najef.2024.102232

Warkulat S, Pelster M (2024) Social media attention and retail investor behavior: Evidence from r/wallstreetbets. *Int Rev Financ Anal* 96: 103721. https://doi.org/10.1016/j.irfa.2024.103721

Wen T, Chen YW, Syed TA, et al. (2025) Examining communication network behaviors, structure and dynamics in an organizational hierarchy: A social network analysis approach. *Inform Process Manag* 62: 103927. https://doi.org/10.1016/j.ipm.2024.103927

Xu Z, Du Z, Cheng Q, et al. (2024) Institutional investor information networks and stock price synchronicity. *Emerg Mark Financ Trade* 60: 3559–3573. https://doi.org/10.1080/1540496X.2024.2354803

Ying J (2024) Gradual information diffusion across commonly owned firms. *J Financ Econ* 156: 103852. https://doi.org/10.1016/j.jfineco.2024.103852

Yu M, Hu X, Zhong A (2024) Network centrality, information diffusion and asset pricing. *Int Rev Financ Anal* 93: 103223. https://doi.org/10.1016/j.irfa.2024.103223

Yu X, Nie Y, Li W, et al. (2024) Key nodes of misinformation source inference: A message-passing-based approach. *Int J Mod Phys C* 35: 2450146. https://doi.org/10.1142/S0129183124501468

Zhai X, Ma H, Zhang Y, et al. (2024) Institutional investor network and idiosyncratic volatility of stocks. *Econ Polit* 36: 1261–1288. https://doi.org/10.1111/ecpo.12289

Zhang Z, Luo Q (2024) Retail investor attention and analyst earnings forecasts: Evidence from China. *Pac-Basin Financ J* 83: 102238. https://doi.org/10.1016/j.pacfin.2023.102238

Zhao X, Zhang Y, Zhai Q, et al. (2024) A multi-attribute decision-making approach for critical node identification in complex networks. *Entropy* 26: 1075. https://doi.org/10.3390/e26121075

Zhu M, Liang Y, Wu Z (2020) Internet information interaction networks and stock price crash risk: Public opinion supervision or irrational contagion. *China Ind Econ* 2020: 81–99. https://doi.org/10.19581/j.cnki.ciejournal.2020.10.003



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)