

Investors' Information Interaction, Sentiment Convergence and Stock Price Co-jumps

Fei Ren

School of Business, East China University of Science and Technology,
No.130 Meilong Road, Shanghai 200237, China
E-mail: fren@ecust.edu.cn
Tel: +86-15821081457

Peng-Cheng Wang

School of Business, East China University of Science and Technology,
No.130 Meilong Road, Shanghai 200237, China
E-mail: Y10220165@mail.ecust.edu.cn

Zhang-HangJian Chen

School of Economics, Anhui University,
No.111 Jiulong Road, Hefei 230601, China
E-mail: hangjian@ahu.edu.cn

Hafiz Hoque*

School of Management, Swansea University,
Singleton Park, Swansea, Wales SA2 8PP, United Kingdom
E-mail: h.hoque@swansea.ac.uk
ORCID: 0000-0002-4354-3895

*Corresponding author

Hafiz Hoque

School of Management, Swansea University,
Singleton Park, Swansea, Wales SA2 8PP, United Kingdom
E-mail: h.hoque@swansea.ac.uk
Tel: +44 (0) 1792 987272
ORCID: 0000-0002-4354-3895

ACKNOWLEDGEMENTS

This work was partially supported by the National Natural Science Foundation (No. 72201003).

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Investors' Information Interaction, Sentiment Convergence and Stock Price Co-jumps

Abstract

Stock price co-jumps play a crucial role in assessing market risk and determining risk premiums, yet research on their drivers remains limited, primarily focusing on the release of public information. Leveraging large-scale textual data from the Eastmoney.com stock forum during 2020-2023, this study demonstrates that individual investors' information interactions significantly influence stock price co-jumps. To establish causality, we use the Eastmoney.com stock forum outage on November 12, 2021 as an exogenous reduction in investors' information interactions and examine its effect on co-jumps. Further analysis shows that positive information interactions exert a stronger effect on upward co-jumps than negative interactions do on downward co-jumps. Mechanism analysis reveals that cross-stock sentiment convergence mediates the relationship between information interactions and co-jump occurrences. Heterogeneity analyses reveal that information interactions exert stronger facilitating effects on stock price co-jumps in pairs of small-cap stocks and in pairs characterized by low information transparency than in other stock pairs. Furthermore, the asymmetric effect of positive versus negative information interactions on upward versus downward co-jumps is significant only for stock pairs in high-tech industries. Finally, based on these empirical findings, we develop an investment strategy that outperforms the Shanghai Stock Exchange Composite Index, achieving an annualized return of 19.41%. This study has important implications for investment strategy formulation and regulatory policy design.

Keywords: Investor information interaction; Sentiment convergence; Co-jumps; Textual analysis; Investment strategy

1. Introduction

In financial markets, stock price co-jumps are defined as significant and synchronized price movements across multiple stocks within extremely short time intervals. Studies have shown that during financial crises, the frequency of such synchronized jumps across financial assets tends to increase (Bollerslev et al., 2008). These co-jumps have profound implications for financial markets and investor behavior, affecting key areas such as asset pricing, portfolio management, and financial risk management (Gilder et al., 2014; Caporin et al., 2017; Ding et al., 2024). Existing literature primarily attributes co-jumps to macroeconomic news, geopolitical events, and firms' earnings announcements (Bollerslev et al., 2008; Bouri et al., 2022; Christensen et al., 2025). Despite these findings, the channels through which information spreads among related stocks and the extent to which such dissemination influences stock price co-jumps remain insufficiently understood. Leveraging post-reply data from Eastmoney.com, this study addresses this gap by examining how individual investors' information interactions on social media platforms, acting as a channel of information transmission among different stocks, influence stock price co-jumps and identifying the underlying mechanisms driving this relationship.

Social media platforms have become crucial channels for investors to obtain and share stock-related information (Cookson et al., 2024; Cao et al., 2025), and investors' information interactions on these platforms facilitate the diffusion of information across stocks (Jiang et al., 2019; Lu & Zhao, 2024). Given the constraints of limited attention, investors typically focus on a small subset of stocks that share fundamental characteristics (Peng & Xiong, 2006). Co-investors who engage in discussions about these related stocks facilitate the dissemination of similar information across them. Specifically, investors focusing on a particular stock form distinct communities within social networks, and co-investors serve as bridges connecting these otherwise isolated communities across relevant stock forums, thereby promoting the spread of shared information through posting, and replying. According to the social contagion model, information-sharing behaviors within each stock-related community can trigger chain reactions, enabling the rapid and widespread diffusion of similar information disseminated by co-investors. This process fosters synchronized trading behaviors across different stocks, leading to price co-movement (Chen et al., 2024). Since co-jumps represent an extreme form of co-movement, it is reasonable to infer that individual investors' information interactions on social media serve as a key driver of co-jumps.

Furthermore, two facets of the existing literature provide insights into the potential mechanisms underlying the impact of individual investors' information interactions on stock price co-jumps. On the one hand, individual investors' information interactions, posts and

replies in different stocks' sub-forums, can establish the information dissemination channels across stocks. These channels expose individual investors to homogeneous information (Chen et al., 2024), thereby triggering similar sentiment responses and ultimately leading to cross-stock sentiment convergence (Huang & Meng, 2024). On the other hand, prior studies have demonstrated that individual investor sentiment can exacerbate the volatility of stock prices (Liu et al., 2022; Yang & Xiong, 2024; Yang et al., 2024). More specifically, the information disseminated on social media platforms, such as policies and technological innovations, has the potential to influence sentiment of individual investors, which leads to single-stock price jumps. By integrating these two strands of research, the individual investors' information interactions initially trigger cross-stock sentiment convergence, which potentially leads to simultaneous price jumps in the same direction. Consequently, we propose that cross-stock sentiment convergence may serve as a micro-mechanism through which individual investors' information interactions impact stock price co-jumps.

Prior studies have demonstrated that positive and negative information have asymmetric impacts on upward and downward jumps in individual stock prices (Stambaugh et al., 2012; Yang & Xiong, 2024). However, there is a lack of research examining the differential effects of positive versus negative information interactions on upward and downward co-jumps, which are defined as simultaneous upward or downward jumps in paired stocks (Gilder et al., 2014; Ding et al., 2024). Given that investors respond asymmetrically to positive and negative information (Zhang & Semmler, 2009; Stambaugh et al., 2012; Yang & Xiong, 2024), positive information is typically associated with expected returns and thus tends to be disseminated more quickly (Mondria et al., 2021; Reichenbach & Walther, 2025), triggering herd behavior (Dimic et al., 2018). In contrast, negative information tends to elicit more cautious responses and spreads more slowly, with investors not necessarily selling their holdings immediately upon receipt (Ding et al., 2004). Therefore, by investigating how positive information interactions drive upward co-jumps and how negative information influences downward co-jumps, this study aims to address the gap in the existing literature regarding the asymmetric effects of positive and negative information interactions on co-jump dynamics.

In this paper, by utilizing the post-reply data during 2020-2023 from Eastmoney.com, China's largest online financial platform, we explore the relationship between individual investors' information interactions and stock price co-jumps. We quantify the daily intensities of information interaction between two stocks by calculating the number of posts and replies on the related sub-forums generated by co-investors, defined as investors who actively participating in both associated sub-forums. Meanwhile, we leverage minute-level high-

frequency data from the RESSET Database¹ to calculate the daily frequency of co-jumps among stocks, following Ding et al. (2024). We then construct a two-way fixed-effects panel data regression model, and examine whether the information interaction intensity has significant positive impacts on stock price co-jumps in the current and subsequent periods, controlling for a series of firm-level and market-level control variables. We identify the causal effect of investor information interactions by exploiting the November 12, 2021 outage of the Eastmoney.com stock forum, which exogenously curtailed communication among individual investors. We also perform a series of robustness checks to ensure the robustness of our findings.

Subsequently, we classify the co-jumps into upward and downward co-jumps based on whether these two stock price jumps are upward jumps or downward jumps, simultaneously. To align with this classification, we measure positive (negative) information interaction intensities by calculating the number of co-investors' posts and replies with positive (negative) sentiment in the two sub-forums. Then, we examine the impacts of positive (negative) information interactions on upward (downward) co-jumps, respectively. The results suggest that positive information interactions have a more significant effect on upward co-jumps compared with the effect of negative information interactions on downward co-jumps. This asymmetry effect has also been corroborated in the existing literature on jumps (Stambaugh et al., 2012; Yang & Xiong, 2024). Heterogeneity tests further reveal that the promoting effect of information interactions on stock price co-jumps is stronger for small-cap stock pairs and for pairs with low information transparency than for other pairs; moreover, the asymmetric impact of positive versus negative information interactions on upward versus downward co-jumps is significant only among high-tech industry pairs.

Moreover, we delve into the micro-mechanism underpinning the impact of individual investors' information interactions on stock price co-jumps, scrutinizing cross-stock sentiment convergence. We measure the daily individual investor sentiment for each stock based on the textual sentiment of posts and replies on the message board, and then calculate the consistency between them as the proxy indicator of cross-stock sentiment convergence. The results of mechanism analysis demonstrate that the cross-stock information diffusion, driven by individual investors' information interactions on social media, can initially trigger the cross-stock sentiment convergence, and subsequently induce stock price co-jumps.

Finally, we construct a dynamic trading strategy based on the predictive capacity of

¹ The RESSET Database: <https://db.resset.com/common/main.jsp>

information interactions for co-jumps. Specifically, we divide the 2020-2023 dataset into eight consecutive six-month windows and examine the impacts of positive (negative) information interactions on upward(downward) co-jumps within each window. We establish trading rules for the subsequent window by leveraging the substantial influence of individual investors' information interactions on co-jumps observed in the preceding window. Specifically, we initiate long positions in stocks with the strongest positive information interactions and short positions in those with the strongest negative interactions before the market closes on each trading day, and close out these positions the following day. With this strategy, we achieve a cumulative return of 69.63% from July 2020 to December 2023, after accounting for transaction costs, significantly outperforming the SSE's cumulative return over the same period (-3.89%), with a difference of 73.52%.

Our research contributes to several strands of literature. To the best of our knowledge, this study represents the inaugural effort to demonstrate that individual investors' information interactions on social media serve as one of the key drivers of co-jumps. In doing so, it supplements existing studies focused on the drivers of co-jumps, while previous research merely explained co-jumps from the perspective of possible common shocks that stocks might face at the macro or industry level (Caporin et al., 2017; Baker et al., 2021; Christensen et al., 2025). This study delves into the specific process of information dissemination on social media platforms. We explore how individual investors' information interactions, particularly within stock-specific sub-forums, facilitate the cross-stock diffusion of information, ultimately triggering co-jumps. This in-depth analysis provides new insights into how information on social media influences stock price co-jumps.

Furthermore, this study adds to the existing literature on stock price co-jumps (Gilder et al., 2014; Ding et al., 2024) by distinguishing co-jumps in different directions and exploring the differences in their drivers. This research also contributes to previous studies that analyze the asymmetric effects of positive and negative sentiments on stock prices (Zhang & Semmler, 2009; Stambaugh et al., 2012; Yang & Xiong, 2024; Yi et al., 2024), by investigating the heterogeneous impacts of positive versus negative sentiment information diffusion on upward and downward co-jumps.

Additionally, our paper extends and enriches the literature focusing on the role of sentiment in stock markets. The existing literature that treats investor sentiment as a direct driver of stock prices (Liu et al., 2022; Jeon et al., 2022; Yang & Xiong, 2024). Instead, we focus on the stock-pair level. By verifying that cross-stock sentiment convergence serves as a key mechanism through which information interactions drive co-jumps, we also enrich the literature on sentiment convergence.

Finally, this study offers valuable insights into the intersection of textual analysis, sentiment convergence, and co-jump risk management. By applying textual analysis to large-scale text data, it extends the measurement of sentiment convergence across stock pairs. Furthermore, we employ textual analysis to examine the asymmetric effects of positive and negative information interactions on upward and downward co-jumps and demonstrate that this asymmetry can be leveraged to generate excess returns. This interdisciplinary approach not only advances research on sentiment convergence measurement but also provides practical implications for co-jump risk management.

The remainder of this paper is structured as follows. Section 2 introduces the sample data and the key variables used in this study. Section 3 presents the empirical models employed in this study. Section 4 presents the empirical findings and discussion. Conclusions are in Section 5.

2. Literature Review and Hypothesis Development

2.1. Individual Investors' Information Interactions and Stock Price Co-jumps

The Chinese stock market is characterized by a dominant presence of individual investors, whose accounts represent over 99% of all market participants. These investors are highly active and account for more than 80% of daily trading volume (Chen et al., 2024). They primarily share and exchange stock-related information on social media platforms such as Eastmoney.com (Jiang et al., 2019; Cao et al., 2025), which features dedicated discussion forums for each listed company. The content of these discussions reflects individual investors' interpretations of public information, including macroeconomic indicators, monetary policies, geopolitical events, and company announcements, and have been empirically linked to co-jumps in stock prices (Bollerslev et al., 2008; Caporin et al., 2017; Baker et al., 2021; Bouri et al., 2022). Importantly, beyond reacting to public information, investor discourse on such platforms also encompasses forward-looking views on market trends, trading strategies, and perceived market dynamics. Therefore, analyzing user-generated posts and replies on Eastmoney.com enables a deeper understanding of how information diffusion among individual investors facilitates the spread of both factual and sentiment-driven content, potentially triggering synchronous price jumps across interrelated stocks.

The theory of limited attention suggests that investors typically focus on a limited set of stocks that share fundamental similarities (Peng & Xiong, 2006). Co-investors who engage in discussions about these related stocks facilitate the dissemination of similar information across them (Chen et al., 2024). Specifically, drawing on narrative economics (Shiller, 2017), we

argue that investors within a specific stock sub-forum form distinct communities centered around particular stocks, and co-investors serve as conduits for cross-stock narrative propagation. Posts and replies shared by co-investors in sub-forums of related stock pairs may contain similar information about these focused stocks due to their fundamental similarities. Once such information reaches each stock sub-forum, investor information interactions within each community trigger cascading effects, causing the information to spread rapidly among a great many investors, in accordance with the social contagion model (Pond et al., 2020). Subsequently, this cross-stock information diffusion may lead a large number of individual investors in these stocks to make trading decisions in the same direction (Yang & Xiong, 2024), ultimately contributing to price co-jumps among them. Moreover, the ongoing diffusion process allows the information to continue propagating and amplifying over time, potentially resulting in a lasting impact of co-investor-shared information on the co-jumps of related stocks. Based on this analysis, we propose the following hypothesis:

H1: Individual investors' information interactions on social media significantly promote both contemporaneous and future stock price co-jumps.

2.2. The Asymmetry Effects on Co-jumps: Positive VS Negative Information Interactions

Following prospect theory (Kahneman & Tversky, 1979), individuals are more sensitive to losses than to equivalent gains, and thus reluctant to realize losses (Ding et al., 2004). Existing research indicates that investors respond asymmetrically to positive and negative information (Zhang & Semmler, 2009; Stambaugh et al., 2012; Yang & Xiong, 2024). Positive information is generally associated with favorable return expectations, leading to faster acceptance and broader dissemination, thereby accelerating cross-stock information diffusion (Josefy et al., 2023). This process stimulates herding behavior on social media platforms, promoting the convergence of optimistic sentiment and shared expectations among investors, which in turn facilitates synchronized buying decisions (Dimic et al., 2018). The resulting surge in buy orders exerts upward pressure on stock prices and may contribute to frequent upward co-jumps. In contrast, negative information is often perceived with skepticism regarding its reliability, resulting in slower propagation across investor networks. Furthermore, when confronted with adverse news, investors are more inclined to hold or even accumulate positions rather than liquidate holdings, a behavioral tendency that delays the price incorporation of negative information (Ding et al., 2004). As a result, the impact of negative information on downward price movements tends to be attenuated. Moreover, in the Chinese A-share market, where short-sale constraints limit the expression of negative sentiment, investors with positive sentiment tend to dominate price formation, rendering stock prices more reflective of positive

expectations (You & Wu, 2012; Yi et al., 2024). Overall, investors on social media exhibit asymmetric responses, favoring positive over negative information. The heightened salience and accelerated diffusion of positive content increase its likelihood of triggering herd behavior, thereby enhancing the probability of frequent upward co-jumps in stock prices. Therefore, we propose the following hypothesis:

H2: Positive information interactions lead to a higher frequency of upward co-jumps relative to downward co-jumps induced by negative information interactions.

2.3. Mechanisms of Investors' Information Interaction and Stock Price Co-jumps

Investor sentiment, as a key driver of trading behavior, has been shown to play a critical role in stock price jumps (Gao et al., 2021; Liu et al., 2022; Gao et al., 2022; Karampatsas et al., 2023). Prior studies document that when investor sentiment within a stock forum becomes aligned, it often leads to a surge in synchronized trading behavior, which subsequently triggers price jumps in that stock (Yang & Xiong, 2024). In the case of two stocks, according to narrative economics (Shiller, 2017), co-investors who actively post or reply in these two sub-forums make similar information publicly accessible to all participants. Contagion theory further posits that these co-investors transmit their sentiments through posts and replies, and other investors in the corresponding sub-forums perceive and internalize such sentiments (Yi et al., 2024). This process drives sentiment convergence—the phenomenon wherein investors in related stocks exhibit similar sentiments—across these stocks. Moreover, Zeng et al. (2019) and Gao & Zhao (2023) suggest that sentiment convergence can prompt investors of related stocks to make synchronized buying or selling decisions, thereby triggering co-jumps among these stocks, although this mechanism has not been empirically validated. Therefore, cross-stock sentiment convergence is likely the key mechanism through which individual investors' information interactions drive stock price co-jumps. Accordingly, we propose the following hypothesis:

H3: Cross-stock sentiment convergence represents a key mechanism by which information interactions among individual investors facilitate co-jumps.

3. Data and Methodology

3.1. Data and Sample Selection

This study selects A-share listed companies on the Shanghai Stock Exchange as the research sample, covering the period from January 1, 2020, to December 31, 2023. To ensure data quality and consistency, several filtering criteria were applied: (1) non-special treatment

(ST) stocks, (2) non-financial stocks, (3) positive undistributed profit, (4) at least 200 trading days per year, and (5) with a price of no less than RMB 3 per share². After screening, 1,144 stocks were retained for analysis. We collect the textual data of individual investors' posts and replies in stock forums corresponding to our sample stocks from Eastmoney.com. We then acquire minute-level high-frequency stock price data for the same sample from the RESSET Database³. In addition, we collect data for control variables from CNRDS⁴, including firm-specific data (e.g., industry, details of institutional shareholding), market-specific data (e.g., market value, trading volume, turnover rate), and information-related data (e.g., analyst reports and news tone).

While our study focuses on information interactions among individual investors on Eastmoney.com, it is important to note that the platform's user base consists predominantly of individual investors from China. In the United States, investor interactions on platforms such as Stocktwits may be shaped by a more mature market structure, a higher proportion of institutional participation, and a more open regulatory environment (Cookson et al., 2024), potentially leading to distinct behavioral dynamics. Consequently, although the findings of this study provide valuable insights into the information-sharing behavior of Chinese individual investors, the generalizability of these results may be constrained when applied to markets with differing investor compositions and institutional frameworks, particularly in the U.S. and other developed economies.

3.2. Measure of Key Variables

3.2.1. Co-jumps

We employ 5-minute stock price data to identify daily co-jumps and calculate the daily number of co-jumps among stock pairs (CJ) using the methodology proposed by Ding et al. (2024). We then classify co-jumps into upward and downward categories and compute the daily number of upward and downward co-jumps, respectively. Specifically, when stocks i and j both exhibit upward jumps simultaneously, this is defined as an upward co-jump (up_CJ); conversely, when both stocks experience downward jumps at the same time, it is defined as a downward co-jump ($down_CJ$). The methodological details are provided in Appendix A.

² In the A-share market, the minimum tick size for order declarations is 0.01 yuan. When a stock's price is too low, this fixed tick size can lead to an over-identification of co-jumps.

³ The RESSET Database: <https://db.resset.com/common/main.jsp>

⁴ Chinese Research Data Services Platform: <https://www.cnrds.com/>

In addition, there is another type of co-jump, known as a mixed-direction co-jump, in which one stock experiences an upward jump while the other experiences a downward jump. The uncertainty regarding which stock undergoes an upward or downward movement leads to high complexity in sentiment divergence between the stock pair, thereby complicating the analysis of the underlying mechanisms of sentiment convergence in co-jumps and increasing the challenges in implementing investment strategies due to heightened decision-making uncertainty. Therefore, this study focuses on upward and downward co-jumps to streamline the analysis, enhance model operability, and provide clearer guidance for investment strategies.

3.2.2. Intensity of Individual Investors' Information Interaction

To measure individual investors' information interactions, we utilize post-reply data from Eastmoney.com⁵, China's largest online financial platform, spanning January 2020 to December 2023. The platform hosts over 40 million monthly active users and generates more than 500,000 posts daily⁶. For each listed company, Eastmoney.com provides a dedicated message board, referred to as a sub-forum, where individual investors exchange information, express opinions, and discuss stock-specific developments.

According to the limited attention theory, investors typically focus on a limit set of stocks that share common fundamental characteristics, such as belonging to the same industry sector or being associated with a similar investment theme (Peng & Xiong, 2006). Within online platforms, investors exchange information and opinions by posting and replying in sub-forums dedicated to these related stocks. Investors who actively engage in multiple sub-forums, referred to as co-investors, serve as conduits for the diffusion of similar information across stocks (Chen et al., 2024). This dissemination of shared information may lead to synchronized trading behavior among investors, ultimately resulting in price co-movements among related stocks. Following the method proposed by Chen et al. (2024), we quantify the daily intensity of individual investors' information interaction between two stocks as the total number of posts and replies generated by co-investors in both sub-forums on a given day, as follows:

$$num_{i,j,t} = \sum_{a=1}^{cu_{i,j,t}} (n_{a,i,t} + n_{a,j,t}), \quad (1)$$

⁵The posts and replies on the Eastmoney.com are generally used for textual analysis in studies on investor sentiment and investor information interaction in the Chinese stock market (Dong et al., 2024; Chen et al., 2024).

⁶ Data Source: <https://www.analysys.cn/>

where $cu_{i,j,t}$ denotes the count of co-investors simultaneously participating in both sub-forums i and j on day t , while $n_{a,i,t}$ and $n_{a,j,t}$ represent the number of posts and replies generated by co-investors in each of the two stock sub-forums on day t .

To examine the effects of positive and negative information interactions by individual investors on upward and downward co-jumps, respectively, we first classify posts and replies as positive or negative based on their sentiment scores, with detailed method provided in Appendix B. Then, we calculate the positive information interaction intensity $num_{i,j,t}^+$ and the negative information interaction intensity $num_{i,j,t}^-$ between sub-forum i and sub-forum j on a given day t as follows:

$$num_{i,j,t}^{+/-} = \sum_{a=1}^{cu_{i,j,t}} n_{a,i,t}^{+/-} + n_{a,j,t}^{+/-}, \quad (2)$$

where $cu_{i,j,t}$ denotes the count of co-investors simultaneously participating in both sub-forum i and sub-forum j on day t , $n_{a,i,t}^{+/-}$ and $n_{a,j,t}^{+/-}$ represent the number of posts and replies generated by co-investors in each of the two stock sub-forums on day t , and “+/-” means that the sentiment of the post/reply is positive or negative, respectively.

3.2.3. Cross-stock Sentiment Convergence

Information disseminated by co-investors across the two sub-forums may gradually spread to the broader investor base in sub-forum, ultimately shaping the sentiment of all participants in both forums. Accordingly, we construct measures of cross-stock sentiment convergence based on the sentiment of all posts and replies in each stock sub-forum. We first calculate a sentiment index for each stock sub-forum using the following procedure: (i) We compute the sentiment scores of posts and replies in each stock sub-forum using the method described in Appendix B and classify them as either positive or negative; (ii) We count the number of positive (negative) posts and replies for each stock on a given day; (iii) Following the approach of Yi et al. (2024), we construct a daily investor sentiment index STM_{it} for stock i by aggregating the counts of positive and negative posts and replies in each stock sub-forum on each trading day, which is defined as follows:

$$STM_{it} = \frac{pos_{it} - neg_{it}}{pos_{it} + neg_{it}}, \quad (3)$$

where pos_{it} represents the total counts of positive posts and corresponding replies generated in the sub-forum of stock i on day t and neg_{it} represents the total counts of negative posts and their associated replies within the same sub-forum for stock i on day t . When the number of

positive posts in the sub-forum is greater than the number of negative posts, STM_{it} is greater than 0 and vice versa is less than 0.

Subsequently, we construct a proxy measure for sentiment convergence following Ton et al. (2024). We construct the investor sentiment consistency indicator SC between stock pairs as the proxy measure for cross-stock sentiment convergence:

$$SC_{i,j,t} = |STM_{it} - STM_{jt}|, \quad (4)$$

where STM_{it} is the stock forum sentiment index of sub-forum i on a given day t , this indicator measures the difference in daily investor sentiment between two stocks. A larger value of $SC_{i,j,t}$ indicates a lower degree of sentiment convergence, whereas a smaller value indicates a higher degree of sentiment convergence.

3.2.4. Control Variables

To ensure the robustness of the research results, we control for established determinants of stock price co-movements documented in the existing literature. Specifically, we selected the following control variables that may impact stock price co-jumps: stock price difference (Chen et al., 2024), trading volume difference (Yi et al., 2024), turnover rate difference (Lee & Bhaskaran, 2000), market value difference (Kumar & Lee, 2006), industry difference (Blau et al., 2023), common analyst coverage (Chen et al., 2024), common institutional ownership (Anton & Polk, 2014), and news tone difference (Jeon et al., 2022). All control variables are measured daily at the stock-pair level. The specific measurement methods are shown in Table1.

[Insert Table 1 Here]

3.3. Regression Model

3.3.1. Individual Investors' Information Interactions and Co-jumps

In order to test the effect of individual investors' information interactions across stocks on price co-jumps, we construct a two-way fixed effects panel data regression model with robust standard errors clustered at the stock-pair level:

$$CJ_{i,j,t} = \alpha_1 + \beta_1 \cdot num_{i,j,t} + \sum \gamma_1 Controls_{i,j,t} + \lambda_i + \lambda_j + \tau_t + \varepsilon_{i,j,t}, \quad (5)$$

where $CJ_{i,j,t}$ is the number of co-jumps among stocks i and j on trading day t , $num_{i,j,t}$ is the intensity of individual investors' information interactions across stock sub-forums related to stocks i and j on that day, and $Controls_{i,j,t}$ includes a set of control variables: stock price difference, trading volume difference, turnover rate difference, market value difference,

industry difference, common analyst coverage, common institutional ownership, and news tone difference. $\lambda_i, \lambda_j, \tau_t, \varepsilon_{i,j,t}$ are the individual fixed effects of stock i , the individual fixed effects of stock j , the time fixed effects, and the random error term, respectively.

3.3.2. Modelling Positive (Negative) Individual Investors' Information Interactions on Upward (Downward) Co-Jumps

In this section, we investigate the effect of positive (negative) individual investors' information interactions on upward (downward) co-jumps. To achieve this, we employ two-way fixed effects panel data regression models with robust standard errors clustered at the stock-pair level, as presented below:

$$up_CJ_{i,j,t} = \alpha_2 + \beta_2 num_{i,j,t}^+ + \sum \gamma_2 Controls_{i,j,t} + \lambda_i + \lambda_j + \tau_t + \varepsilon_{i,j,t}, \quad (6)$$

$$down_CJ_{i,j,t} = \alpha_3 + \beta_3 num_{i,j,t}^- + \sum \gamma_3 Controls_{i,j,t} + \lambda_i + \lambda_j + \tau_t + \varepsilon_{i,j,t}, \quad (7)$$

where $up_CJ_{i,j,t}$ is the number of upward co-jumps between stocks i and j on trading day t , and $down_CJ_{i,j,t}$ is the number of downward co-jumps between stocks i and j on that day. $num_{i,j,t}^+$ and $num_{i,j,t}^-$ are the intensities of positive and negative individual investors' information interactions across stock sub-forums related to stocks i and j on that day.

3.3.3. Mechanism Analysis

To examine whether sentiment convergence serves as the mechanism through which individual investors' information interactions drive stock price co-jumps, we conduct a two-stage analysis following the methodology employed in recent studies (Safiullah et al., 2024). In the first stage, we estimate the effect of individual investors' information interactions on stock price co-jumps using model (5). In the second stage, we employ models (8) to assess the impact of individual investors' information interactions on the sentiment convergence. The model specifications are as follows:

$$SC_{i,j,t} = \alpha_4 + \beta_4 \cdot num_{i,j,t} + \sum \gamma_4 Controls_{i,j,t} + \lambda_i + \lambda_j + \tau_t + \varepsilon_{i,j,t}, \quad (8)$$

where $SC_{i,j,t}$ is the proxy indicator of sentiment convergence between stocks i and j on trading day t . $num_{i,j,t}$ is the intensity of individual investors' information interactions across stock sub-forums related to stocks i and j on that day.

Furthermore, we also distinguish between positive and negative information interactions to investigate whether both types facilitate sentiment convergence, and whether their effects differ in magnitude or significance. The model specifications are as follows:

$$SC_{i,j,t} = \alpha_5 + \beta_5 \cdot num_{i,j,t}^+ / num_{i,j,t}^- + \sum \gamma_5 Controls_{i,j,t} + \lambda_i + \lambda_j + \tau_t + \varepsilon_{i,j,t} \quad (9)$$

where $SC_{i,j,t}$ is the proxy indicator of sentiment convergence between stocks i and j on trading day t . $num_{i,j,t}^+$ and $num_{i,j,t}^-$ are the intensities of positive and negative individual investors' information interactions across stock sub-forums related to stocks i and j on that day.

Additionally, following Ullah & Begum (2025), we employ the Sobel-Goodman test and the Bootstrap method to quantify the indirect effect and calculate the proportion of the total effect mediated by sentiment convergence.

4. Empirical Results

4.1. Description of Key Variables

Table 2 presents the descriptive statistics for the key variables utilized in this paper, including the dependent variables, independent variables, and control variables. From January 1, 2020, to December 31, 2023, the total number of downward co-jumps (11,006,991) is greater than that of upward co-jumps (8,219,525). This observation aligns with findings from prior studies, which indicate that downward jumps occur more frequently than upward jumps (Bollerslev et al., 2008). The mean of individual investors' information interaction intensity num is 1.412 with a standard deviation of 6.549, suggesting that significant information interactions exist in certain stock pairs. Moreover, the mean values of positive individual investors' information interaction intensity num^+ and negative individual investors' information interaction intensity num^- are 0.704 and 0.537, respectively, indicating that positive information interactions are more frequent than negative interactions in the sample. The higher intensity of positive individual investors' information interactions, relative to negative information interactions, is likely attributable to investors' greater propensity to accept and disseminate positive information (Josefy et al., 2023). In addition, the mean of SC is 0.406, suggesting the presence of sentiment convergence across different stock pairs. These findings confirm the existence of both individual investors' information interactions and sentiment convergence in the sample, providing a solid foundation for the subsequent regression analysis. Given the excess zeros in dependent variables, we employ Poisson pseudo-maximum likelihood method as a robustness check to ensure the reliability of our results. To examine the relationships among these key variables, we compute the Pearson correlation coefficients for the main variables, as presented in Appendix C.

[Insert Table 2 Here]

4.2. Individual Investors' Information Interactions and Stock Price Co-Jumps

To explore the influence of individual investors' information interactions across stock sub-forums on stock price co-jumps, we use the intensity of individual investors' interactions as the explanatory variable and the number of stock co-jumps as the dependent variable. Table 3 presents the baseline regression results assessing the relationship between individual investors' information interactions and stock price co-jumps. Across all models (columns 1–6), the coefficients of num_t remain consistently positive and statistically significant at the 1% level. This result provides strong evidence that individual investors' information interactions across stock sub-forums significantly contribute to stock price co-jumps, thereby supporting our Hypothesis H1. Moreover, the robustness of these results is confirmed through various model specifications. After including control variables, the coefficient of num_t in column 2 remains positive and significant. In addition, when the dependent variables are shifted forward up to four periods ($CJ_{t+1} \sim CJ_{t+4}$, columns 3 to 6), the coefficients remain statistically significant, indicating that individual investors' information interactions also influence future stock price co-jumps. These findings are consistent with prior research highlighting the role of individual investors' information interactions in driving stock price co-movements (Chen et al., 2024; Jiang et al., 2019). While prior research focused on information diffusion and its effects on stock co-movements, our study extends this literature by specifically identifying how information exchanges across stock sub-forums contribute to stock price co-jumps.

[Insert Table 3 Here]

4.3. Robustness Tests and Endogeneity Issues

4.3.1. Robustness Tests

(1) Robustness Tests: Poisson pseudo-maximum likelihood (PPML) method

The dependent variable $CJ_{i,j,t}$ is non-negative integer, which may violate the normality assumption of the baseline linear model. To address this, we re-estimate our main specifications using PPML method, a natural framework for count data. Given the long sample period (2020–2023) and to ensure that our results are not driven by time-varying unobserved factors or structural breaks, we estimate the Poisson model separately for each calendar year (2020, 2021, 2022, 2023). This year-by-year approach also serves as an additional robustness check against potential over-fitting or cross-temporal heterogeneity.

[Insert Table 4 Here]

(2) Robustness Tests: Alternative Co-jumps Measure

We transform the dependent variable into a binary indicator num_t^* , assigning a value of 1

if a co-jump occurred between the stock pairs, and 0 otherwise. The regression results for this specification are presented in column 1 of Table 5. The findings indicate that the intensity of individual investors' information interactions continues to have a statistically significant effect on the occurrence of co-jumps, thus confirming the robustness of our original results.

(3) Robustness Tests: Excluding Extreme Values

To ensure the robustness of the baseline regression results, we conduct a sensitivity test to assess the influence of extreme values. Specifically, we exclude the top 5% of observations with the highest num_t and re-estimate using the remaining subsample. As shown in column 2 of table 5, the coefficient on num_t remains largely unchanged in magnitude and statistical significance compared to the baseline estimate from the full sample. This indicates that the main findings of the study are not driven by a small number of highly active information interactions.

(4) Robustness Tests: Normalizing the information interaction intensity

To mitigate the potential influence of variations in shareholder base size across listed firms on individual investors' information interaction intensities, we redefine the measure of information interaction intensity by normalizing it with respect to the total number of shareholders in each period, resulting in a "per capita information interaction intensity" (N_num_t). Regression results based on this adjusted metric remain highly consistent with our baseline findings, as shown in columns 3 and 4 of table 5, indicating that the baseline results are robust to differences in shareholder scale and are not driven by firm-level variations in shareholder counts.

(5) Robustness Tests: Extreme Event Analysis

In this robustness check, we focus on the extreme event associated with the Chinese government's reduction of the stamp duty on August 28, 2023, which had a substantial impact on the stock market. Such extreme events can introduce large and unexpected shocks to market dynamics. We conduct this analysis to demonstrate that our results remain robust even in the context of such extreme events. To evaluate whether our results remain valid during this period, we limit the sample to August and September 2023. The results, reported in column 5 of Table 5, indicate that even during this period of heightened market volatility, the effect of individual investors' information interactions on stock price co-jumps remains statistically significant, further confirming the robustness of our conclusions.

(6) Robustness Tests: Annual Subsample Analysis

In addition, we divide the full sample from 2020 to 2023 into four annual subsamples and

conduct separate regressions for each. This division allows us to examine whether the relationship between individual investors' information interactions and stock price co-jumps remains consistent across different market conditions. By utilizing annual subsamples, we can ensure that our results are not driven by specific years or temporal effects. The results presented in columns 6 to 9 of Table 5 demonstrate that the relationship between individual investors' information interaction intensity and co-jumps remains significant and consistent across all four years, further confirming the impact of information interactions on stock price co-jumps.

Notably, the COVID-19 pandemic began in 2020, and the Russia-Ukraine war started in 2022. Despite these exogenous shocks, the coefficient for num_t remains significantly positive across all years, indicating that even when accounting for these macroeconomic events, the impact of information interactions on stock price co-jumps remains robust.

[Insert Table 5 Here]

4.3.2. Endogeneity Issues

In the preceding regression analyses, we have already controlled for potential common shocks, such as news tone differences, common institutional ownership, common analyst coverage, and have tested the impact of current information interactions on co-jumps in the following 1 to 4 periods. However, despite these efforts, reverse causality remains a potential source of endogeneity. Specifically, stock price co-jumps may also drive increased individual investors' interactions, as significant price movements often prompt more discussions and information exchanges across stock sub-forums. To address this concern, we employ exogenous shock and instrumental variables approaches to mitigate the bias arising from reverse causality.

(1) Endogeneity: Exogenous Shock: Eastmoney.com Forum Outage

To address potential endogeneity concerns and strengthen the causal link between individual investors' information interactions and stock price co-jumps, we utilize the Eastmoney.com stock forum outage on November 12, 2021 as an exogenous shock that substantially reduces individual investors' information interactions. On that day, the Eastmoney.com forum suffered an outage starting at around 7:00, before the opening of the A-share market. The service was not restored until the market closed, leaving investors unable to access the platform during the entire trading session. Representatives from Eastmoney.com subsequently confirmed the outage, stating that maintenance work was underway, and services were fully restored later in the day. The unexpected technical failure served as an external shock that temporarily reduced the intensity of information exchanges among individual investors,

independent of the concurrent stock price movements. This exogenous event provides a unique opportunity to examine the impact of investor information interactions on stock price co-jumps, while addressing the concern of reverse causality.

Specifically, we adopt a difference-in-differences (DID) approach to mitigate potential endogeneity issues related to unobservable stock pair-level characteristics. Since the forum outage only lasted one day and all functions resumed normal operation shortly afterwards, its impact is transient. We therefore define our sample as two weeks before the outage and two consecutive trading days covering the outage day and the next trading day. We first construct a time indicator variable, *Post*, which equals 1 for these two post-event trading days, and 0 for the two weeks preceding the outage. Next, we classify stock pairs with non-zero information interaction intensity before the outage as the treatment group ($Treat=1$), and those with no information interactions before the outage as the control group ($Treat=0$). We then replace the baseline information interaction measure in our co-jump specification with the interaction term $Treat \times Post$ to estimate the causal effect of the forced reduction in forum activity on co-jump behavior.

As shown in column (1) of Table 6, the coefficient of $Treat \times Post$ is -0.0013 and statistically significant, which suggests that the Eastmoney forum outage significantly reduced the frequency of stock price co-jumps for treatment pairs relative to control pairs. This finding confirms that the exogenous shock-induced reduction in investor information interactions weakens the co-jump behavior, consistent with our main hypothesis that individual investors' information interactions drive stock price co-jumps. This DID design mitigates reverse causality concerns, providing robust causal evidence of the positive relationship between investor information interactions and stock price co-jumps.

[Insert Table 6 Here]

(2) Endogeneity: Instrumental Variables Approach

To further address endogeneity issues, we apply the instrumental variable approach. We construct the first instrumental variable based on the total information interaction volume of all other stock portfolios and simultaneously control for the interaction fixed effects of industry and time to alleviate possible industry spillover problems. Then, we construct the second instrumental variable using the internal instrumental variable method based on heteroskedasticity proposed by Lewbel (2012). The results continue to show a significant effect of information interaction intensity on stock price co-jumps. Detailed results are provided in Appendix D.

4.4. Further Analysis

4.4.1. The Asymmetry of Driving Factors for Upward and Downward Co-jumps

Table 7 presents the regression results for positive (negative) individual investors' information interactions on upward (downward) co-Jumps, with the sample period spanning from January 2020 to December 2023.

In Panel A, across columns 1-9, the coefficients of num_t^+ consistently show a positive and statistically significant relationship with upward co-jumps. Specifically, in the column 1 (2020.01–2023.12), the coefficient of num_t^+ is 0.0002, indicating that a higher level of positive information interaction significantly increases the likelihood of upward co-jumps among stock pairs. The relationship remains significant across the sub-periods, with the coefficients ranging from 0.0001 to 0.0003. In Panel B, we observe the impact of negative information interaction num_t^- on downward co-jumps $down_CJ_t$. Similar to Panel A, we find statistically significant positive coefficients for num_t^- across most periods. In the column 1 (2020.01–2023.12), the coefficient for num_t^- is 0.0001, with a t-statistic of 20.35, indicating a positive impact on downward co-jumps, and this effect is consistent across most sub-periods.

We compare the regression results reported in Panel A and Panel B and find a clear asymmetry in the effects of positive versus negative information interactions on stock price co-jumps. Positive information interactions, denoted as num_t^+ , exhibit a stronger and more immediate influence on upward co-jumps, as indicated by the larger and more statistically significant coefficients in Panel A. In contrast, negative information interactions, denoted as num_t^- , demonstrate a weaker effect on downward co-jumps, as reflected in Panel B, where the coefficients are smaller and less statistically significant. This asymmetry in the impact of positive and negative information interactions on stock price co-jumps may stem from differential investor responses to each type of information. Positive information tends to elicit faster and more synchronized reactions from investors, thereby triggering upward co-jumps. Negative information, on the other hand, often disseminates more slowly and may be accompanied by greater investor skepticism, resulting in a less pronounced effect on downward co-jumps. These findings provide empirical support for Hypothesis 3, which suggests that positive information interactions have a stronger influence on upward co-jumps than negative information interactions have on downward co-jumps.

These results are based on the Chinese market, where individual investors dominate and may exhibit a bias toward positive information. This asymmetry in the investor base

may contribute to the observed behavioral patterns, as individual investors are often more susceptible to sentiment-driven decision-making. However, it is important to acknowledge that in institutional-heavy markets, such as those in the United States, institutional investors are likely to adopt a more rational and data-driven approach, which could lead to different outcomes.

[Insert Table 7 Here]

4.4.2. Heterogeneity Analysis

(1) Market Capitalization Heterogeneity

To examine whether the effect of individual investors' information interaction intensity on stock price co-jumps varies across firms with different market capitalizations, we perform a heterogeneity analysis by categorizing stock pairs based on their market capitalization. We divide stock pairs into two groups following Dong et al. (2024): "Big-Cap" if both stocks in the pair have a market capitalization exceeding the median size, and "Small-Cap" if both stocks have a market capitalization below the median size. The dynamics of information diffusion and the sensitivity to investor sentiment may differ between "Big-Cap" and "Small-Cap" stocks. "Big-Cap" stocks typically have more stable investor bases and lower volatility, making them less susceptible to co-jumps driven by information dissemination. In contrast, "Small-Cap" stocks tend to be more volatile and more responsive to investor sentiment and information flows, which may render them more prone to co-jumps triggered by information interactions.

The results of the market capitalization heterogeneity analysis are presented in Table 8 Panel A, where columns (1)-(2) correspond to the estimates for the "Big-Cap" stock pairs and columns (3)-(4) to those for the "Small-Cap" stock pairs. As shown, the coefficients for the "Small-Cap" stock pairs are not only statistically more significant but also exhibit larger magnitudes. This suggests that information interactions among individual investors have a stronger influence on co-jumps in the case of "Small-Cap" stock pairs. Additionally, we find that in the "Small-cap" samples, the impact of positive information interaction on upward jumps is much greater than that of negative information interaction on downward jumps.

(2) Industry Heterogeneity

Subsequently, building on the approach of Hoque & Mu (2023), we analyze industry-level heterogeneity to evaluate whether the observed asymmetry is more pronounced in specific thematic sectors. This examination is motivated by evidence that co-investors, particularly individual investors, tend to concentrate their investments within certain industries. Following

the method proposed by Peng & Mao (2017), and in accordance with the *Guidelines on Industry Classification of Listed Companies* issued by the China Securities Regulatory Commission (CSRC) in 2012, we classify the sample into two subsamples: the “High-tech” stock pairs and the “Others” stock pairs, based on whether both stocks in each stock pair belong to the high-tech industry.

The empirical results reported in columns (1) and (2) of Table 8 Panel B indicate that, in the “High-tech” stock pairs, the effect of positive information interactions on upward jumps is significantly stronger than that of negative information interactions on downward jumps. In contrast, this asymmetry is not statistically significant in the “Others” stock pairs. The results can be interpreted from two perspectives: on the one hand, the high-tech industry contains a relatively high proportion of small-cap stocks, thereby inheriting the strong sensitivity of small-cap stocks to positive information interactions, which gives rise to a pronounced positive-negative asymmetry. On the other hand, the high-tech industry itself is characterized by rapid technological change, high information asymmetry, and strong dependence on innovation-related news and market expectations. Individual investors in this sector tend to interpret optimistic information more positively, further amplifying the driving effect of positive information interactions.

(3) Information Transparency Heterogeneity

To examine whether the impact of individual investor information interaction on stock price co-jumps varies among enterprises with different levels of information transparency, we group the stock pairs based on the financial report transparency and conduct a heterogeneity analysis. Following the research of Hutton et al. (2009), we construct an information transparency index, which is obtained by taking the average of the absolute values of the discretionary accruals of the enterprise over the past three years. The larger the index value, the lower the information transparency of the enterprise. Following the grouping logic of the previous heterogeneity analysis, the stock pairs are divided into two groups: if the information transparency index of both stocks in each stock pairs is below the median, they are classified as the “high transparency” stock pairs, indicating higher information transparency; if the opacity index of both stocks in each stock pairs is above the median, they are classified as the “low transparency” stock pairs, indicating lower information transparency.

The results of the information transparency heterogeneity analysis are presented in Table 8 Panel C. As shown, for the “low transparency” stock pairs, the coefficients of both num_t^+ and num_t^- are larger and more statistically significant than those for the “high transparency” stock pairs. This suggests that information interactions among individual

investors have a stronger influence on co jumps for firms with lower information transparency. This is because stocks with low information transparency disclose insufficient public financial information and business conditions, making it difficult for investors to form accurate judgments based on public information. In this situation, individual investors rely more on information interaction on social media platforms such as stock forums to obtain incremental information and make up for the insufficiency of public information. Therefore, information interaction plays a stronger marginal role in low transparency stocks and has a more significant driving effect on stock price co-jumps. Given that small-cap stocks are typically characterized by lower information transparency and fewer analyst followings (Lang et al., 1996), the above mechanism is particularly relevant for explaining the stronger effects observed in the small-cap segment.

[Insert Table 8 Here]

4.4.3. Mechanism Analysis

(1) Mechanism Analysis: Two-stage Analysis

To empirically test whether sentiment convergence serves as the mechanism through which individual investors' information interactions drive stock price co-jumps, we conduct a two-stage analysis based on the method used in recent studies (Safiullah et al., 2024), with panel data regressions with fixed time and individual effects, clustered at the stock pair level. In the first stage, we examine the effect of individual investors' information interactions on stock price co-jumps. In the second stage, we test whether information interactions influence sentiment convergence, which in turn may impact stock price co-jumps. We employed two-way fixed effects in our model to address major concerns of omitted variables, thereby supporting a causal interpretation of the mediation pathway.

Table 9 presents the results of the mechanism analysis. In column (1), the coefficient of num_t is positive and statistically significant, suggesting that increased information interactions among individual investors leads to a higher probability of stock price co-jumps. In column (2), the coefficient of SC_t is negative and statistically significant. This suggests that higher levels of individual investors' investor information interaction reduce sentiment divergence between stocks, further confirming that individual investors' information interactions contribute to sentiment convergence across different stock sub-forums.

We further examine whether positive and negative information interactions also operate through sentiment convergence. As shown in columns (3) and (4), the coefficient

of num_t^+ on upward co-jumps is positive and significant, while its coefficient on SC_t is negative and significant. This indicates that positive information interactions promote upward co-jumps by reducing sentiment divergence. Similarly, columns (5) and (6) show that num_t^- has a positive effect on downward co-jumps and a negative effect on SC_t , suggesting that negative interactions also facilitate downward co-jumps via sentiment convergence. Collectively, these findings provide robust evidence for the hypothesis that individual investors' information interactions drive stock price co-jumps through the mechanism of sentiment convergence.

[Insert Table 9 Here]

(2) Mechanism Analysis: Sobel-Goodman test and Bootstrap

Additionally, to provide more robust evidence of the mediation mechanism, we employ the Sobel-Goodman test and the Bootstrap method. These methods are standard in causal mediation analysis and directly test the significance of the indirect effect (Ullah & Begum, 2025), thereby establishing a causal link between information interactions and co-jumps through sentiment convergence. The results, presented in Table 10, are consistent with our previous conclusions. The mediation analysis reveals that sentiment convergence explains approximately 3.4% of the total effect of overall information interactions on co-jumps. For positive information interactions on upward co-jumps, the mediated proportion is 1.4%; for negative information interactions on downward co-jumps, it is 5.6%. These proportions, while modest, are statistically significant, confirming that sentiment convergence serves as a partial mediator across all specifications.

[Insert Table 10 Here]

4.4.4. Investment strategies

According to the empirical results presented in Sections 4.4.1 and 4.4.2, positive (negative) information interactions among individual investors significantly contribute to the occurrence of upward (downward) co-jumps, and the effects exhibit heterogeneous characteristics across firms of different sizes. Furthermore, using a two-way fixed-effects panel regression model with time and stock fixed effects, we evaluate the predictive power of co-jumps on future individual stock returns, as detailed in Appendix E. Based on these findings, we develop a dynamic trading strategy using a rolling-window approach. The dataset covering the period from 2020 to 2023 is divided into eight consecutive six-month windows. Within this framework, historical data are used to determine out-of-sample trading strategies, with daily trading

decisions in the current window informed by the influence of positive and negative information interactions on co-jumps observed in the previous window. Moreover, the holding time for each position is dynamically adjusted based on the predictive power of co-jumps for future returns, which is also estimated from the results of the preceding window. The formal introduction of margin trading and short selling in China on March 31, 2010, and the subsequent expansion of the program to include over 1,000 stocks, provides the foundation for constructing and analyzing both long and short strategies in this study. The implementation details of the trading strategy are provided in Appendix F.

We conduct out-of-sample back-tests of the strategies involving long and short positions on “Big-Big” and “Small-Small” stock pairs using data spanning from July 1, 2020, to December 31, 2023. In the performance evaluation, the transaction costs in the Chinese A-share market are accounted for and appropriately adjusted⁷. The performance is benchmarked against the SSE Composite Index, with evaluation focusing on key metrics including cumulative return, Sharpe ratio, Sortino ratio, max drawdown, VaR 95%, and CVaR 95%, and tracking error (Kim et al., 2018), as shown in Table 11. We also present the cumulative returns of each strategy during the testing period in Fig. 1 and find that the “Long S-S” strategy demonstrates consistent outperformance relative to the SSE Index across the entire back-testing period, with cumulative returns of 69.63%.

[Insert Table 11 and Fig. 1 Here]

5. Conclusion

This study explores the impact of individual investors’ information interactions on stock price co-jumps, with a particular focus on the role of social media platforms as a channel for information diffusion. We find compelling evidence that information interactions across different stock sub-forums significantly contribute to stock price co-jumps. Then, we observe an asymmetry in the effects of positive and negative interactions, with positive individual investors’ information interactions having a stronger and more consistent impact on upward co-jumps, while negative individual investors’ information interactions exert a weaker influence on downward co-jumps. Furthermore, the mechanism

⁷ For long positions, the total costs include brokerage commissions at 0.015% of the trade value, a unilateral securities transaction stamp tax of 0.05% levied on sellers as per State Taxation Administration regulations, share transfer fees at 0.001%, and ancillary charges such as exchange fees (0.00487%). Short positions incur additional financing costs beyond the baseline fees, most notably margin interest charged at an annualized rate of 7%.

analysis reveals that sentiment convergence is the mechanism through which individual investors' information interactions drive stock price co-jumps. Moreover, the heterogeneity analysis reveals that the promoting effects of information interactions on co-jumps are stronger for small-cap stocks, and firms with low information transparency. Notably, the asymmetric impact of positive versus negative information interactions on upward versus downward co-jumps is uniquely significant in high-tech sectors, whereas this asymmetry is not observed in other industries. Additionally, building on these findings, we propose a dynamic investment strategy based on information interactions and co-jumps. Our strategy, which adjusts positions according to the occurrence of co-jumps and utilizes predictive capacity for future returns, demonstrates robust performance in back-tests, with substantial returns outperforming the SSE Index.

This study makes several important contributions to the literature. First, it advances research on stock price co-jumps by highlighting the pivotal role of the cross-stock of information diffusion driven by individual investors' information interactions on social media platforms, offering new insights into the drivers of co-jump dynamics. Second, it extends the literature on investor sentiment in financial markets by shifting the focus from individual stocks to stock pairs and identifying cross-stock sentiment convergence as the key mechanism through which information interactions drive co-jumps. Thirdly, this study enriches the understanding of co-jumps by distinguishing between upward and downward co-jumps and examining the asymmetric effects of positive versus negative information interactions among individual investors. Finally, it offers valuable insights into the intersection of textual analysis, sentiment convergence, and co-jumps, not only advancing the measurement of the sentiment convergence index but also delivering practical implications for co-jump risk management.

Our findings offer practical implications for both investors and regulators. For investors, this study provides valuable insights into how information interactions among individual investors on social media platforms influence stock price co-jumps. Individual investors can leverage these interactions to improve their trading strategies, particularly in small-cap stocks that are more sensitive to such dynamics. For regulators, the research underscores the importance of monitoring and regulating information dissemination on social media. As individual investors increasingly rely on these platforms for financial information, the spread of both accurate and misleading content can significantly affect market stability. Regulators should consider implementing measures to promote timely and transparent disclosure, counter misinformation, and monitor online forums for signs of market manipulation, thereby helping to mitigate risks associated with excessive co-

movement in stock prices.

This study has two limitations that point to promising avenues for future research. First, while we examine the effects of positive and negative information interactions on upward and downward co-jumps, our analysis does not account for their impact on mixed-direction co-jumps. Future studies could investigate this dimension to achieve a more comprehensive understanding of how information interactions influence co-jump dynamics, as well as to refine investment strategies accordingly. Second, our analysis is based on the Chinese market, where individual investors dominate and may exhibit a preference for positive information. To extend the scope of this research, future work could compare markets with differing investor compositions, such as the individual investor-driven Chinese market and the institutionally dominated U.S. market, to examine how information interactions affect stock price co-jumps in distinct institutional and behavioral environments.

References

- Anton, M., & Polk, C. (2014). Connected stocks. *The Journal of Finance*, 69(3), 1099-1127.
- Baker, S. R., Bloom, N., Davis, S. J. & Sammon, M. C. (2021). What Triggers Stock Market Jumps? Working Paper at <https://doi.org/10.3386/w28687>.
- Blau, B. M., Griffith, T. G., & Whitby, R. J. (2023). Industry regulation and the comovement of stock returns. *Journal of Empirical Finance*, 73, 206-219.
- Bollerslev, T., Law, T. H., & Tauchen, G. (2008). Risk, jumps, and diversification. *Journal of Econometrics*, 144(1), 234-256.
- Bouri, E., Gupta, R., & Vo, X. V. (2022). Jumps in geopolitical risk and the cryptocurrency market: the singularity of Bitcoin. *Defence and Peace Economics*, 33(2), 150-161.
- Cao, J., Zou, Z., Yang, X., & Xiong, X. (2025). How does social media shape stock liquidity? The role of investor online communication networks. *International Review of Financial Analysis*, 103, 104236.
- Caporin, M., Kolokolov, A., & Renò, R. (2017). Systemic co-jumps. *Journal of Financial Economics*, 126(3), 563-591.
- Chen, Z. H., Wu, W. L., Li, S. P., Bao, K., & Koedijk, K. G. (2024). Social media information diffusion and excess stock returns co-movement. *International Review of Financial Analysis*, 91, 103036.
- Christensen, K., Timmermann, A., & Veliyev, B. (2025). Warp speed price moves: Jumps after earnings announcements. *Journal of Financial Economics*, 167, 104010.
- Cookson, J. A., Lu, R., Mullins, W. & Niessner, M. (2024). The social signal. *Journal of Financial Economics* 158, 103870.
- Dimic, N., Neudl, M., Orlov, V., & Äijö, J. (2018). Investor sentiment, soccer games and stock returns. *Research in International Business and Finance*, 43, 90-98.
- Ding, D. K., Charoenwong, C., & Seetoh, R. (2004). Prospect theory, analyst forecasts, and stock returns. *Journal of Multinational Financial Management*, 14(4-5), 425-442.
- Ding, Y., Li, Y., Liu, G., Zheng, X. (2024). Stock co-jump networks. *Journal of Econometrics*, 239(2), 105420.
- Dong, L., Yu, B., Qin, Z., & Lam, K. S. (2024). Liquidity risk and expected returns in China's stock market: A multidimensional liquidity approach. *Research in International Business and Finance*, 69, 102247.
- Eierle, B., Klammer, S., & Muck, M. (2022). Does it really pay off for investors to consider information from social media? *International Review of Financial Analysis*, 81, 102074.
- Gao, Y., Zhao, C., Sun, B., & Zhao, W. (2022). Effects of investor sentiment on stock volatility: new evidences from multi-source data in China's green stock markets. *Financial Innovation*, 8(1), 77.
- Gao, X., Gu, C., & Koedijk, K. (2021). Institutional investor sentiment and aggregate stock returns. *European Financial Management*, 27(5), 899-924.
- Gao, Y., & Zhao, C. (2023). Investor sentiment and stock price jumps: A network analysis based on China's carbon-neutral sectors. *The North American Journal of Economics and Finance*, 68, 101954.
- Gilder, D., Shackleton, M. B., & Taylor, S. J. (2014). Cojumps in stock prices: Empirical evidence. *Journal of Banking & Finance*, 40, 443-459.

- Hoque, H., & Mu, S. (2023). Information spillover in Chinese hybrid IPO auctions. *Journal of International Money and Finance*, 131, 102767.
- Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R2, and crash risk. *Journal of financial Economics*, 94(1), 67-86.
- Huang, X., & Meng, Y. (2024). Information shock, market reaction, and stock message board information diffusion. *The Quarterly Review of Economics and Finance*, 95, 180-192.
- Jeon, Y., McCurdy, T.H., Zhao, X. (2022). News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Journal of Financial Economics* 145, 1–17.
- Jiang, L., Liu, J., & Yang, B. (2019). Communication and comovement: Evidence from online stock forums. *Financial Management*, 48(3), 805-847.
- Josefy, A. B., Rees, L., & Tse, S. (2023). Investor response to bad news versus good news management earnings guidance. *Accounting Horizons*, 37(1), 121-142.
- Karampatsas, N., Malekpour, S., Mason, A., & Mavis, C. P. (2023). Twitter investor sentiment and corporate earnings announcements. *European financial management*, 29(3), 953-986.
- Kim, J. H., Kim, W. C., Kwon, D. G., & Fabozzi, F. J. (2018). Robust equity portfolio performance. *Annals of Operations Research*, 266(1), 293-312.
- Kumar, A., & Lee, C. M. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, 61(5), 2451-2486.
- Lee, C. M., & Swaminathan, B. (2000). Price momentum and trading volume. *the Journal of Finance*, 55(5), 2017-2069.
- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of business & economic statistics*, 30(1), 67-80.
- Liu, W., Zhang, C., Qiao, G., & Xu, L. (2022). Impact of network investor sentiment and news arrival on jumps. *The North American Journal of Economics and Finance*, 62, 101780.
- Lu, S., & Zhao, J. (2024). Investor network and stock return comovement: Information-seeking through intragroup and intergroup followings. *International Review of Financial Analysis*, 93, 103204.
- Mondria, J., Wang, X., & Wu, T. (2021). Familiarity and Surprises in International Financial Markets: Bad news travels like wildfire; good news travels slow. *Journal of International Money and Finance*, 115, 102390.
- Patatoukas, P. N. (2012). Customer-base concentration: Implications for firm performance and capital markets: 2011 American accounting association competitive manuscript award winner. *The accounting review*, 87(2), 363-392.
- Peng, H. M., & Mao, X. S. (2017). Government Subsidies for Innovation, Company Executives Background and R&D Investment-Evidence from the High-tech Industry. *Finance & Trade Economics*, 38(3), 147-161.
- Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3), 563-602.
- Pond, T., Magsarjav, S., South, T., Mitchell, L., & Bagrow, J. P. (2020). Complex contagion features without social reinforcement in a model of social information flow. *Entropy*, 22(3), 265.
- Reichenbach, F., & Walther, M. (2025). Attention allocation of investors on social media: the role of prospect theory. *Journal of Behavioral Finance*, 26(3), 317-334.
- Safiullah M., Phan D., Kabir M. (2024). Green innovation and corporate default risk. *Journal*

- of International Financial Markets, Institutions and Money, 95, 102041.
- Shiller, R. J. (2017). Narrative economics. *American Economic Review*, 107(4), 967-1004.
- Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2), 288-302.
- Ton, T., Leung, H., Gao, Y., & Schiereck, D. (2024). Agreeing to disagree: Informativeness of sentiments in internet message boards. *Pacific-Basin Finance Journal*, 87, 102485.
- Ullah, S., & Begum, M. (2025). FinTech and financial sustainability: A mediating role of financial inclusion. *Technological Forecasting and Social Change*, 217, 124129.
- Yang, N., Fernandez-Perez, A., & Indriawan, I. (2024). Spillover between investor sentiment and volatility: The role of social media. *International Review of Financial Analysis*, 96, 103643.
- Yang, J., & Xiong, Y. (2024). Social media sentiment contagion and stock price jumps and crashes. *Pacific-Basin Finance Journal*, 88, 102520.
- Yi, M. M., Ren, F., He L. J., & Xiong, X. (2024). Research on the Impact of Sentiment Contagion on Stock Price Co-movement: Evidence from Big Data of Stock Forum. *Journal of Management Science*. 37(04), 147-160. [In Chinese]
- You, J. X., & Wu J. (2012). Spiral of silence: media sentiment and the asset mispricing. *Economic Research Journal*, 47(7), 141–152. [In Chinese]
- Lang, M. H., & Lundholm, R. J. (1996). Corporate disclosure policy and analyst behavior. *Accounting review*, 71(4), 467-492.
- Zeng, Q., Liu, S., & Zhang, Q. (2019). Common sentiment and price contagion. *Journal of Systems Science and Complexity*, 32(5), 1426-1437.
- Zhang, W., & Semmler, W. (2009). Prospect theory for stock markets: Empirical evidence with time-series data. *Journal of Economic Behavior & Organization*, 72(3), 835-849.

Table 1. List of Variables

Notation	Meaning	Calculation Method
CJ	the number of co-jumps without distinction of direction	See Appendix A.
up_CJ	the number of upward co-jumps	
$down_CJ$	the number of downward co-jumps	
num	information interaction intensity	See formula (12).
$num^{+/-}$	positive/negative information interaction intensity	See formula (13).
SC	the investor sentiment consistency between stock pairs	See formula (15).
Dp	stock price difference	$ \ln(p_i/p_j) $, where p_i represents the price of stock i .
$Dvol$	trading volume difference	$ \ln(vol_i/vol_j) $, where vol_i represents the trading volume of stock i .
$Dsize$	market value difference	$ \ln(size_i/size_j) $, where $size_i$ represents the market value of stock i .
Dhs	turnover rate difference	$ \ln(hs_i/hs_j) $, where hs_i represents the turnover rate of stock i .
$DInd$	industry difference	1 for the same industry, 0 for otherwise
$same_holders$	common institutional ownership	It represents whether there is common institutional shareholding.
$common_analysts$	common analyst coverage	It represents whether there is common analyst coverage.
$Dnews_tone$	news tone difference	$ news_tone_i - news_tone_j $, where $news_tone_i$ is calculated by subtracting the number of negative news articles from the number of positive news articles related to the stock i .

Note: This table presents the main variables and their calculation methods. We standardize all variables to daily frequency. Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, and $Dnews_tone$ are daily measures. Since individual stock institutional holdings are disclosed quarterly, the common institutional ownership variable ($same_holders$) adopts quarterly values and stays unchanged across all days in each quarter. The $common_analysts$ is measured based on all analyst reports issued during each month. We then apply this monthly value to all daily observations in the corresponding month.

Table 2. Description of Variables

Variables	Mean	Std. Dev.	Min	Max
<i>CJ</i>	0.048	0.218	0	6
<i>up_CJ</i>	0.015	0.121	0	4
<i>down_CJ</i>	0.020	0.139	0	3
<i>num</i>	1.412	6.549	0	5047
<i>num</i> ⁺	0.704	2.862	0	2473
<i>num</i> ⁻	0.537	3.048	0	2040
<i>SC</i>	0.406	0.346	0	2
<i>Dp</i>	0.910	0.746	0	6.908
<i>Dvol</i>	1.809	1.446	0	12.989
<i>Dsize</i>	1.222	0.975	0	8.169
<i>Dhs</i>	1.036	0.804	0	8.867
<i>DInd</i>	0.033	0.178	0	1
<i>same_holders</i>	0.291	0.454	0	1
<i>common_analysts</i>	0.004	0.064	0	1
<i>Dnews_tone</i>	0.162	0.889	0	207

Table 3. Baseline Results

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	CJ_t	CJ_t	CJ_{t+1}	CJ_{t+2}	CJ_{t+3}	CJ_{t+4}
num_t	0.0003*** (30.48)	0.0002*** (28.41)	0.0003*** (28.53)	0.0002*** (26.23)	0.0001*** (24.64)	0.0001*** (23.81)
Dp		-0.0020*** (-38.71)	-0.0018*** (-36.32)	-0.0018*** (-35.63)	-0.0018*** (-35.00)	-0.0018*** (-34.83)
$Dvol$		-0.0007*** (-32.86)	-0.0007*** (-34.06)	-0.0007*** (-33.99)	-0.0007*** (-34.89)	-0.0007*** (-36.25)
$Dsize$		-0.0006*** (-13.32)	-0.0004*** (-8.41)	-0.0005*** (-11.46)	-0.0005*** (-12.16)	-0.0006*** (-13.65)
Dhs		-0.0003*** (-9.71)	0.0002*** (6.88)	-0.0001*** (-4.63)	-0.0003*** (-11.24)	-0.0004*** (-15.08)
$DInd$		0.0035*** (19.44)	0.0035*** (19.29)	0.0036*** (19.81)	0.0036*** (19.99)	0.0036*** (20.05)
$same_holders$		-0.0000*** (-21.60)	-0.0000*** (-24.58)	-0.0000*** (-21.87)	-0.0000*** (-21.36)	-0.0000*** (-22.56)
$common_analysts$		0.0002*** (2.61)	0.0002*** (2.73)	0.0003*** (3.73)	0.0003*** (4.13)	0.0003*** (4.41)
$Dnews_tone$		0.0002*** (17.28)	0.0002*** (14.22)	0.0001*** (10.37)	0.0001*** (10.37)	0.0002*** (15.75)
Constant	0.0480*** (1769.59)	0.0531*** (628.80)	0.0530*** (624.38)	0.0533*** (641.28)	0.0535*** (647.99)	0.0536*** (650.10)
Stock Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.0791	0.0867	0.0865	0.0862	0.0860	0.0858
N (million)	557.2257	557.2257	555.1547	553.0838	551.0129	548.9420

Note: Table 3 reports the baseline results for the impact of individual investors' information interactions on stock co-jumps. num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , $CJ_{t,...,t+4}$ denote the number of co-jumps on trading day $t, ..., t+4$, respectively. Controls includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017).

Table 4. Results of PPML method

	(1)	(2)	(3)	(4)
	CJ_t (2020)	CJ_t (2021)	CJ_t (2022)	CJ_t (2023)
num_t	0.0022*** (12.93)	0.0016*** (10.31)	0.0016*** (16.63)	0.0134*** (20.67)
Constant	-1.8760*** (-644.34)	-2.5391*** (-852.40)	-2.1824*** (-721.68)	-2.4164*** (-882.58)
Controls	YES	YES	YES	YES
Time Fixed	YES	YES	YES	YES
Stock Fixed	YES	YES	YES	YES
Pseudo R^2	0.2391	0.1235	0.1605	0.1389
N (million)	113.1498	133.1320	151.9654	157.5304

Note: This table presents the robustness test results based on the Poisson Pseudo-Maximum Likelihood (PPML) estimation approach. num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , CJ_t is the number of co-jumps on that day, *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, *same_holders*, *common_analysts*, and *Dnews_tone*. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the z-statistics, computed based on standard errors clustered at the stock-pair level in all specifications.

Table 5. Results of Other Robustness Tests

	(1)	(2)	(3)	(4)	(5)
	CJ_t^*	CJ_t	CJ_t	CJ_{t+1}	CJ_t
num_t	0.0003*** (30.49)	0.0028*** (211.89)			0.0086*** (36.25)
N_num_t			54.5932*** (166.35)	60.8429*** (176.70)	
Constant	0.0474*** (1777.88)	0.0511*** (783.42)	0.0527*** (182.37)	0.05250*** (1837.76)	0.0575*** (318.21)
Controls	YES	YES	YES	YES	YES
Time Fixed	YES	YES	YES	YES	YES
Stock Fixed	YES	YES	YES	YES	YES
R^2	0.0787	0.0797	0.0795	0.0795	0.1360
N (million)	557.2257	529.3644	557.2257	555.1547	28.2609

Table 5. (Continued)

	(6)	(7)	(8)	(9)
	CJ_t (2020)	CJ_t (2021)	CJ_t (2022)	CJ_t (2023)
num_t	0.0003*** (22.52)	0.0001*** (14.89)	0.0001*** (11.37)	0.0035*** (6.95)
Constant	0.0594*** (393.12)	0.0473*** (391.42)	0.0500*** (369.59)	0.0596*** (41.26)
Controls	YES	YES	YES	YES
Time Fixed	YES	YES	YES	YES
Stock Fixed	YES	YES	YES	YES
R^2	0.1471	0.0452	0.0771	0.1069
N (million)	113.1498	133.1320	151.9654	157.5304

Note: This table shows the results of additional robustness tests for our baseline regression model. num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , N_num_t is the normalized individual investors' information interaction intensity across two stock sub-forums on trading day t , CJ_t^* is the binary variable that takes the value of 1 if a co-jump happens on that day and 0 otherwise, CJ_t is the number of co-jumps on that day, $mixed_CJ_t$ is the number of mixed-direction co-jumps on that day, CJ_{t+1} is the number of co-jumps on day $t + 1$, *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017).

Table 6. Exogenous Shock: Eastmoney Forum Outage

	(1) CJ_t
$Treat$	0.0004 (1.52)
$Treat \times Post$	-0.0013** (-2.31)
Controls	YES
Stock/Time Fixed	YES
N (million)	4.3838

Note: This table presents the results of the difference-in-differences (DID) analysis using Eastmoney Forum Outage as an exogenous shock. CJ_t is the number of co-jumps on that day, *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, *same_holders*, *common_analysts*, and *Dnews_tone*. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively.

Table 7. The Asymmetry of Driving Factors for Upward and Downward Co-jumps

Panel A: Upward co-jumps									
	All Sample	2020.01- 2020.06	2020.07- 2020.12	2021.01- 2021.06	2021.07- 2021.12	2022.01- 2022.06	2022.07- 2022.12	2023.01- 2023.06	2023.07- 2023.12
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t	up_CJ_t
num_t^+	0.0002*** (26.79)	0.0003*** (12.98)	0.0004*** (24.75)	0.0003*** (14.18)	0.0002*** (16.98)	0.0001*** (11.18)	0.0002*** (12.12)	0.0029*** (46.26)	0.0031*** (63.63)
R^2	0.0649	0.1289	0.0585	0.0226	0.0192	0.0613	0.0676	0.0352	0.1235
N (million)	557.2257	54.3715	58.7783	64.4559	68.6760	73.3765	78.5889	76.7597	80.8394
Panel B: Downward co-jumps									
	All Sample	2020.01- 2020.06	2020.07- 2020.12	2021.01- 2021.06	2021.07- 2021.12	2022.01- 2022.06	2022.07- 2022.12	2023.01- 2023.06	2023.07- 2023.12
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$	$down_CJ_t$
num_t^-	0.0001*** (20.35)	-0.0002*** (-9.69)	0.0001*** (8.82)	0.0002*** (13.36)	0.0001*** (13.43)	0.0001*** (7.65)	0.0002*** (10.08)	0.0018*** (12.70)	0.0004*** (7.88)
R^2	0.1021	0.2314	0.0717	0.0343	0.0430	0.0845	0.0567	0.0292	0.0339
N (million)	557.2257	54.3715	58.7783	64.4559	68.6760	73.3765	78.5889	76.7597	80.8394
Note: This table shows the asymmetric effect results of directional information interactions on stock co-jumps, with Panel A examining the effect of positive individual investors' information interactions on upward co-jumps and Panel B examining the effect of negative individual investors' information interactions on downward co-jumps. $num_t^{+/-}$ are positive/negative individual investors' information interaction intensities across two stock sub-forums on trading day t , up_CJ_t is the number of upward co-jumps on that day, $down_CJ_t$ is the number of downward co-jumps on that day. <i>Controls</i> includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017).									

Table 8. The results of heterogeneity analysis

Panel A. Market capitalization heterogeneity				
	Large Cap		Small Cap	
	(1) up_CJ_t	(2) $down_CJ_t$	(3) up_CJ_t	(4) $down_CJ_t$
num_t^+	0.0001*** (16.72)		0.0008*** (45.42)	
num_t^-		0.0001*** (16.58)		0.0003*** (22.95)
R^2	0.0802	0.0838	0.0695	0.0900
N (million)	100.5249	100.5249	184.8171	184.8171
Panel B. Industry Heterogeneity				
	High-tech		Others	
	(1) up_CJ_t	(2) $down_CJ_t$	(3) up_CJ_t	(4) $down_CJ_t$
num_t^+	0.0003*** (20.12)		0.0002*** (14.51)	
num_t^-		0.0002*** (12.01)		0.0003*** (14.39)
R^2	0.0746	0.0832	0.0680	0.0861
N (million)	153.2531	153.2531	125.7452	125.7452
Panel C. Information Transparency Heterogeneity				
	high transparency		low transparency	
	(1) up_CJ_t	(2) $down_CJ_t$	(3) up_CJ_t	(4) $down_CJ_t$
num_t^+	0.0001*** (10.66)		0.0004*** (30.88)	
num_t^-		0.0001*** (9.80)		0.0002*** (16.04)
R^2	0.0715	0.0887	0.0710	0.0766
N (million)	2054.9863	2054.9863	86.2467	86.2467

Note: This table reports the comprehensive heterogeneous regression results of individual investors' directional information interactions on stock price co-jumps, with Panel A examining the heterogeneity across market capitalization groups, Panel B across industry types, and Panel C across information transparency levels. $num_t^{+/-}$ are positive/negative individual investors' information interaction intensities across two stock sub-forums on trading day t , up_CJ_t and $down_CJ_t$ are the numbers of upward and downward co-jumps on the same day. Controls includes a set of control variables: Dp , $Dvol$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017).

Table 9. Results of Two-stage Analysis

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	CJ_t	SC_t	up_CJ_t	SC_t	$down_CJ_t$	SC_t
num_t	0.0002*** (28.41)	-0.0011*** (-29.36)				
num_t^+			0.0002*** (26.79)	-0.0021*** (-31.95)		
num_t^-					0.0001*** (20.35)	-0.0019*** (-31.64)
Controls	YES	YES	YES	YES	YES	YES
Stock /Time Fixed	YES	YES	YES	YES	YES	YES
R^2	0.0871	0.0963	0.0649	0.1098	0.1021	0.1099
N (million)	557.2257	504.3567	557.2257	504.3567	557.2257	504.3567

Note: This table presents the results of the two-stage mechanism analysis, which examines the role of sentiment convergence as the underlying channel through which individual investors' information interactions affect stock price co-jumps. SC_t is the sentiment consistency indicator between two sub-forums on the day t . num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , and CJ_t is the number of co-jumps on that day. *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, *same_holders*, *common_analysts*, and *Dnews_tone*. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017).

Table 10. Results of Sobel-Goodman Test and Bootstrap

Panel A Sobel-Goodman test									
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	CJ_t	SC_t	CJ_t	up_CJ_t	SC_t	up_CJ_{tt}	$down_CJ_t$	SC_t	$down_CJ_t$
num_t	0.0002*** (28.41)	-0.0038*** (-385.03)	0.0002*** (26.83)						
SC_t			-0.0022*** (-77.42)						
num_t^+				0.0002*** (26.79)	-0.0021*** (-379.27)	0.0002*** (24.64)			
SC_t						-0.0003*** (-17.50)			
num_t^-							0.0001*** (20.36)	-0.0019*** (-385.97)	0.0001*** (18.70)
SC_t									-0.0019*** (-102.38)
Controls	Yes	YES	YES	Yes	YES	YES	Yes	YES	YES
Stock /Time Fixed	Yes	YES	YES	Yes	YES	YES	Yes	YES	YES
R^2	0.0871	0.0107	0.0945	0.0649	0.0107	0.0734	0.1021	0.0100	0.1063
N (million)	557.2257	504.3567	504.3567	557.2257	504.3567	504.3567	557.2257	504.3567	504.3567
Sobel	0.000*** (z = 77.34)			0.000*** (z = 17.50)			0.000*** (z = 102.16)		
Aroian (Goodman-1)	0.000*** (z = 77.34)			0.000*** (z = 17.50)			0.000*** (z = 102.16)		
Goodman (Goodman-2)	0.000*** (z = 77.34)			0.000*** (z = 17.50)			0.000*** (z = 102.16)		
Indirect effect	0.0000*** (z = 77.34)			0.0000*** (z = 17.50)			0.0000*** (z = 102.16)		
Direct effect	0.0003*** (z = 26.83)			0.0002*** (z = 24.64)			0.0001*** (z = 20.02)		

Total effect	0.0003*** (z = 28.41)			0.0002*** (z = 26.79)			0.0001*** (z = 20.35)		
Proportion of total effect that is mediated	0.034			0.014			0.031		
Panel B Bootstrap method									
	Observed coefficient	z	P-value	Observed coefficient	z	P-value	Observed coefficient	z	P-value
Indirect effect	0.0000	71.92	0.0000	0.0000	16.58	0.0000	0.0000	99.65	0.0000
Direct effect	0.0003	27.62	0.0000	0.0002	25.37	0.0000	0.0001	19.87	0.0000
Note: This table reports mediation effect results from the Sobel-Goodman test and the Bootstrap method. Panel A corresponds to the Sobel-Goodman test and Panel B to the Bootstrap method. SC_t is sentiment consistency indicator between two sub-forums on day t . num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , and CJ_t is the number of co-jumps on that day. <i>Controls</i> includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively.									

Table 11. Performance of Strategies

Strategy	Cumulative Return	Annualized Return	Volatility	Sharpe Ratio	Sortino Ratio
Long B-B	-0.1198	-0.0145	0.0134	-0.2133	-0.3083
Long S-S	0.6963	0.1941	0.0149	0.7072	0.9881
Short B-B	-0.4899	-0.1554	0.0137	-0.8686	-1.2401
Short S-S	-0.3606	-0.0983	0.0141	-0.5864	-0.9859
SSE Index	-0.0389	0.0003	0.0098	-0.1945	-0.2726

Table 13. (Continued)

Strategy	Tracking Error	VaR 95%	CVaR 95%	Max Drawdown
Long B-B	0.0070	0.0220	0.0013	-0.4459
Long S-S	0.0111	0.0250	0.0008	-0.3211
Short B-B	0.0224	0.0224	0.0024	-0.5434
Short S S	0.0220	0.0233	0.0022	-0.5008
SSE Index	0.0000	0.0161	0.0010	-0.2275

Note: This table presents the backtesting performance results of the trading strategies constructed based on our research findings, with the Shanghai Composite Index (SSE Index) included as the benchmark for comparison. Cumulative Return is the return achieved over the entire back testing period. Annualized Return is the compounded return if the strategy were held for a full year. Volatility is measured as the daily standard deviation of returns. Sharpe ratio and Sortino ratio represent the excess return per unit of deviation and downside deviation, respectively, based on daily returns. Max Drawdown shows the maximum peak-to-trough decline over the period. VaR 95% and CVaR 95% indicate the 95th percentile Value at Risk and Conditional Value at Risk, respectively, representing the maximum loss not exceeded with 95% confidence and the expected loss given that the VaR threshold is exceeded. Tracking Error is the standard deviation of the difference between the strategy returns and the benchmark returns (SSE Index). The SSE Index column shows the returns from investing in the Shanghai Composite Index funds, used as a benchmark for comparison.

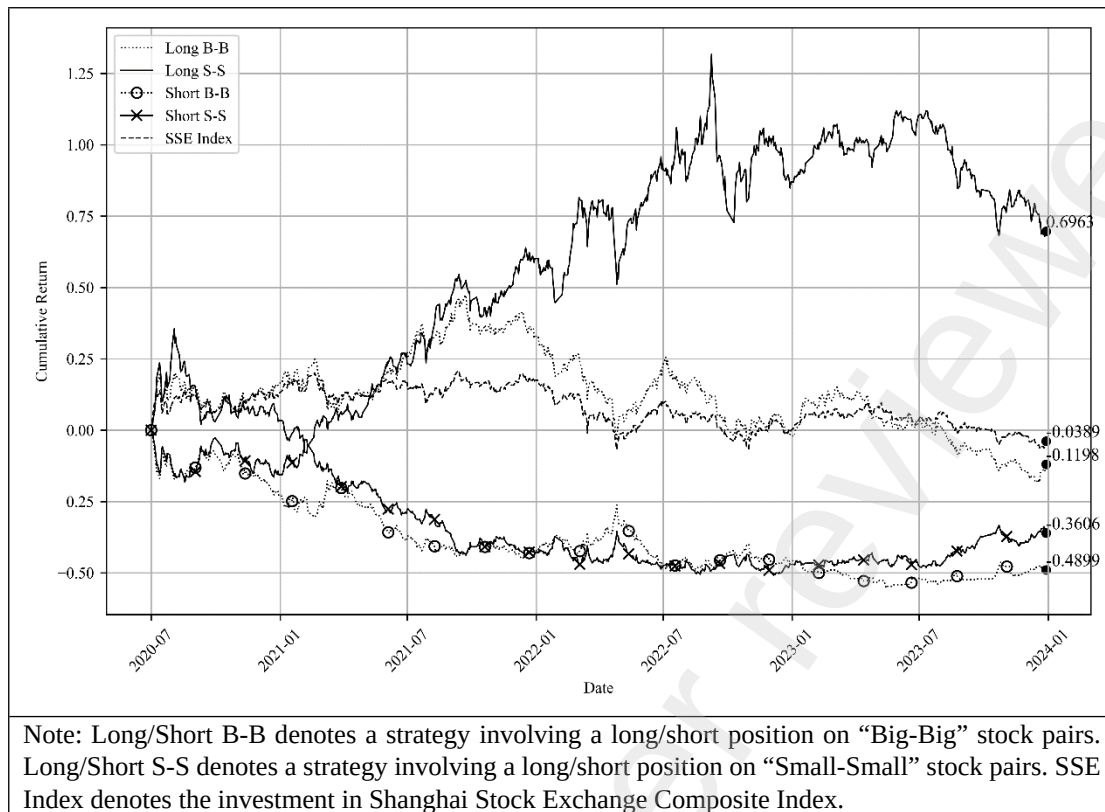


Fig.1 Cumulative Returns of strategies

Online Appendix Summary

Appendix A. Method for Measuring Co-jumps

Appendix B. Method for Measuring the Sentiment of Posts and Replies

Appendix C. Correlation Analysis

Appendix D. Instrumental Variables Approach

Appendix E. Predictive Capacity Test of Co-jumps

Appendix F. Implementation Details of the Strategy

Appendix A. Method for Measuring Co-jumps

Co-jumps

We use 5-minute stock price data to identify daily co-jumps among stock pairs by utilizing the method proposed by Ding et al. (2024), which is detailed as follows:

Step 1. Modeling Log-Price Process and Jump Dynamics. Consider N log-price processes $\{Y_i(t)\}_{i=1}^N$, for each i , the log-price process is an Ito process, the differential form of which is given by:

$$dY_i(t) = \mu_i(t)dt + \sigma_i(t)dB_i(t) + \kappa_i(t)dJ_i(t), \quad (A1)$$

$\mu_i(t)$ represents the drift component of the price, $\sigma_i(t)$ denotes the instantaneous volatility, $B_i(t)$ is the standard Brownian motion, and $(J_i(t): t \geq 0)$ is the Poisson jump process, with $\kappa_i(t)$ being the size of the jump. The log-price change is defined as $\Delta Y_i(t) = Y_i(t) - Y_i(t-1)$, which is non-zero only when a jump occurs in the time interval t .

Step 2. Collecting Intraday Log Return. Assuming that stock prices are sampled at a frequency of M , the s -th intraday logarithmic return of stock i on day t can be denoted as follows:

$$r_{i,t,s} = Y_i\left(t - 1 + \frac{s}{M}\right) - Y_i\left(t - 1 + \frac{s-1}{M}\right), s = 1, 2, \dots, M. \quad (A2)$$

Step 3. Calculation of Realized Volatility and Realized Bipower Variation. The Realized Volatility $RV_{i,t}$ of stock i on trading day t is calculated as follows:

$$RV_{i,t} = \sum_{s=1}^M r_{i,t,s}^2. \quad (A3)$$

As the sampling frequency M tends to infinity, the limit of realized volatility is:

$$\lim_{M \rightarrow \infty} RV_{i,t} \xrightarrow{P} \int_{t-1}^t \sigma_i^2(u)du + \sum_{t-1 < s \leq t} (\Delta Y_i(s))^2, \quad (A4)$$

where $\int_{t-1}^t \sigma_i^2(u)du$ and $\sum_{t-1 < s \leq t} (\Delta Y_i(s))^2$ are the integrated volatility and quadratic variations due to jump components, respectively. Barndorff-Nielsen & Shephard (2004) introduced the realized bipower variation $BV_{i,t}$ as follows:

$$BV_{i,t} = \frac{\pi}{2} \frac{M}{M-1} \sum_{s=1}^{M-1} |r_{i,t,s+1}| |r_{i,t,s}|, \quad (A5)$$

and they proved that $BV_{i,t}$ is consistent with the accumulation volatility as M tends to infinity:

$$\lim_{M \rightarrow \infty} BV_{i,t} \xrightarrow{P} \int_{t-1}^t \sigma_i^2(u) du, \quad (A6)$$

Step 4. Detecting Jumps in Individual Stocks. The daily threshold level $v_{i,t}$ for each stock i is calculated as follows:

$$v_{i,t} = \tau \times \left(\frac{1}{M}\right)^{0.49} \times \sqrt{\min(RV_{i,t}, BV_{i,t})}, \quad (A7)$$

where τ is a tuning parameter, which is set to 4. Jumps can be consistently detected as the sampling frequency M tends to infinity, given that $\mu_i(t)$ and $\sigma_i(t)$ satisfy certain regularity conditions.

For stock i , the total number of jumps detected on a given day t is defined as:

$$J_{i,t} = \sum_{1 \leq s \leq M} 1_{\{|r_{i,t,s}| > v_{i,t}\}}, \quad (A8)$$

where $1_{\{*\}}$ is an indicator function that takes the value of 1 when the condition is met and 0 otherwise.

Step 5. Calculation of Daily Number of Co-jumps. Following the method proposed by Gilder et al. (2014), the co-jumps among two stocks are defined as the intersection of their jumps without distinction of direction. Accordingly, the number of co-jumps between stocks i and j on a trading day t is specified as follows:

$$CJ_{i,j,t} = \sum_{1 \leq s \leq M} 1_{\{|r_{i,t,s}| > v_{i,t}, |r_{j,t,s}| > v_{j,t}\}}. \quad (A9)$$

Step 6. Identifying Co-jumps in Different Directions. In addition, we classify co-jumps into upward and downward categories and calculate the daily number of upward co-jumps and downward co-jumps, respectively. Specifically, when two stocks i and j exhibit upward jumps at the same time, we define this as an upward co-jump (up_CJ). Conversely, when two stocks exhibit downward jumps simultaneously, we define this as a downward co-jump ($down_CJ$). The formulas for computing the daily number of upward and downward co-jumps are as follows:

$$up_CJ_{i,j,t} = \sum_{1 \leq s \leq M} 1_{\{|r_{i,t,s}| > v_{i,t}, |r_{j,t,s}| > v_{j,t}, r_{i,t,s} > 0, r_{j,t,s} > 0\}}, \quad (A10)$$

$$down_CJ_{i,j,t} = \sum_{1 \leq s \leq M} 1_{\{|r_{i,t,s}| > v_{i,t}, |r_{j,t,s}| > v_{j,t}, r_{i,t,s} < 0, r_{j,t,s} < 0\}}. \quad (A11)$$

Appendix B. Method for Measuring the Sentiment of Posts and Replies

We first employ the SnowNLP module in Python, a tool designed for natural language processing of Chinese text, to calculate the sentiment scores of posts and replies based on their textual content. The score represents the probability that the text sentiment is positive. The closer the score is to 1, the more positive the sentiment of the text; conversely, the closer the score is to 0, the more negative the sentiment. Although SnowNLP includes a basic lexicon for positive and negative sentiments, this lexicon is primarily designed for comment data from e-commerce platforms and may not be fully adequate for effectively analyzing the sentiment of financial market texts. To enhance the accuracy of sentiment computation, we have expanded the lexicons by incorporating data from multiple resources: the CNKI (China National Knowledge Infrastructure) sentiment dictionary, the NTUSD Simplified Chinese sentiment dictionary from National Taiwan University, and the BosonNLP sentiment dictionary from the Boson Chinese Semantic Development Platform. After undergoing manual refinement, these resources constitute our final corpus for sentiment analysis.

Then, we use the sentiment score of each post or reply to determine its sentiment as follows: a post/reply with a sentiment score greater than 0.6 is considered positive; a post/reply with a sentiment score greater than 0.4 and less than or equal to 0.6 is considered neutral; and a post/reply with a sentiment score less than or equal to 0.4 is regarded as negative.

To validate the accuracy of the sentiment classification, a random sample of 500 posts and replies is manually coded from the stock forum, and the results generated by our machine-based sentiment classification method are compared against the manual annotations. The overall accuracy rate of our method is 76.1%. Furthermore, the recall rates for positive and negative content are 90.5% and 94.4%, respectively, while the precision rates are 77.7% and 75.8%, respectively. Notably, the high recall values indicate that the classifier effectively identifies most instances of both positive and negative sentiment. These performance metrics are considered acceptable for use in subsequent empirical analysis.

Appendix C. Correlation Analysis

Correlation Analysis

To better understand the relationships between the key variables, we compute the Pearson correlation coefficients among the main variables, presented in Table C. The correlation between individual investors' information interaction intensity num and the number of co-jumps CJ is 0.009, indicating that higher levels of investor information interaction intensity are associated with more frequent co-jumps. Both positive individual investors' information interaction intensity num^+ and negative individual investors' information interaction intensity num^- are positively correlated with upward co-jumps up_CJ and downward co-jumps $down_CJ$, with coefficients of 0.008 and 0.004, respectively. These results suggest that information interactions, irrespective of their sentiment, may contribute to synchronous price jumps. In terms of the relationship with sentiment convergence, the correlation coefficients between num with SC is -0.086, suggesting that the information interactions may facilitate the sentiment convergence across stocks. Additionally, the correlation coefficients between num^+/num^- and SC are -0.085 and -0.081, indicating that both positive and negative information interactions may facilitate sentiment convergence. These findings offer preliminary evidence supporting the role of information interactions in promoting both sentiment convergence and stock price co-jumps.

Table C. Pearson Correlation Coefficient

	CJ	up_CJ	$down_CJ$	num	num^+	num^-	SC
CJ	1						
up_CJ	0.552***	1					
$down_CJ$	0.634***	-0.001***	1				
num	0.009***	0.004***	0.004***	1			
num^+	0.008***	0.005***	0.003***	0.943***	1		
num^-	0.009***	0.003***	0.006***	0.957***	0.821***	1	
SC	-0.004***	0.001***	-0.004***	-0.086***	-0.085***	-0.081***	1

Appendix D. Instrumental Variables Approach

(1) Endogeneity: Instrumental Variables Approach

To further address endogeneity concerns, we employ the instrumental variables approach. Our instrument is constructed as the sum of information interactions across all other stock pairs, excluding the target stock pair i and j . This instrument satisfies the relevance condition. Market-wide information activities generally co-move across stock pairs driven by common factors including overall market sentiment, trading volume and macroeconomic fundamentals, so the constructed instrument is strongly correlated with the information interaction intensity of the target stock pair.

Admittedly, cross-stock information spillovers may exist and threaten the exclusion restriction. We mitigate this concern through two rigorous designs. First, we include “industry $\times$ time” fixed effects to absorb potential industry-specific spillovers and common shocks that could otherwise bias the exclusion restriction. Second, our instrument is aggregated based on a broad sample of 1,144 stocks in the market. Beyond market and industry spillovers, additional information spillovers may arise from firms connected with the target pair via business relationships or capital ties. Nevertheless, given that these related stocks represent only a small share of the total sample (Patatoukas, 2012), their idiosyncratic spillover effects are substantially diluted during summation and exert minimal influence on the co-jump of the target stock pair. Accordingly, our instrument could satisfy the exclusion restriction.

The detailed construction of this instrumental variable is specified as follows:

$$IV1_{i,j,t} = \sum_{\substack{k \neq i \text{ or } j \\ l \neq i \text{ or } j}} \sum_{a=1}^{cu_{k,l,t}} (n_{a,k,t} + n_{a,l,t}), \quad (10)$$

where $cu_{k,l,t}$ is the number of co-investors simultaneously participating in both sub-forum k and sub-forum l on day t , while $n_{a,k,t}$ and $n_{a,l,t}$ represent the total number of posts and replies made by co-investors in each of the two stock sub-forums on day t .

Then, we employ a Two-Stage Least Squares (2SLS) methodology to estimate the causal relationship between individual investors' information interactions and co-jumps, with industry-by-time fixed effects included in both regression stages. The results are reported in the following table D1. In the first stage, we examine how the $IV1_t$ affects the intensity of individual investors' information interactions num_t . As shown in column (1), the coefficient of $IV1_t$ is -0.0013 and is highly statistically significant. In the second stage, we use the

predicted values of individual investors' information interaction intensity from the first stage as the instrumental variable to assess its impact on stock price co-jumps. As presented in column (2), the coefficient of $\widehat{num}_t$ (the predicted num_t) is 0.0010, which is positive and statistically significant. This finding suggests that greater intensity of individual investors' information interactions increases the likelihood of stock price co-jumps, thereby supporting the robustness of our main results.

Table D1. Results of Instrumental Variable Approach ($IV1_t$)

	(1) num_t	(2) CJ_t
$\widehat{num}_t$		0.0010*** (76.05)
$IV1_t$	-0.0013*** (-100.65)	
Controls	YES	YES
Industry×Time Fixed	YES	YES
Stock /Time Fixed	YES	YES
N (million)	557.2257	557.2257
Underidentification test		
Kleibergen-Paap rk LM statistic	8456.80	
P-value	0.00	
Weak identification test		
Cragg-Donald Wald F statistic	1.0e+08	
Kleibergen-Paap rk Wald F statistic	1.0e+04	

Note: This table shows the instrumental variable regression results using $IV1_t$ as the instrument. num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , CJ_t is the number of co-jumps on that day, *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and stock fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively.

To further ensure that the instrumental variables satisfy the exogeneity condition, this paper re-examines the results using the heteroskedasticity-based instrumental variable approach proposed by Lewbel (2012). Lewbel (2012) demonstrates that if the residuals from regressing the endogenous variable on the other exogenous variables in the model exhibit heteroskedasticity, then the product of these residuals and the centered exogenous variables can serve as a valid instrument. Specifically, we generate an internal instrument denoted as $IV2_t$. This approach is particularly useful when conventional instruments are weak or unavailable, as it constructs instruments from the heteroskedastic covariance restrictions in the first-stage residuals, thereby providing a complementary source of identification that does not rely on

external exclusion restrictions. The results are reported in the following table and confirm that our main findings remain robust.

Table D2. Results of Instrumental Variable Approach ($IV2_t$)

	(1) <i>num_t</i>	(2) <i>CJ_t</i>
<i>IV2_t</i>		0.0001*** (10.11)
<i>SC_g</i>	0.1835*** (16.45)	
<i>Dp_g</i>	-0.1433*** (-11.99)	
<i>Dvol_g</i>	0.0456*** (5.63)	
<i>Dsize_g</i>	0.1701*** (16.22)	
<i>Dhs_g</i>	-0.0271** (-2.27)	
<i>DInd_g</i>	0.6393*** (14.62)	
<i>same_holders_g</i>	0.0015*** (8.96)	
<i>common_analysts_g</i>	-0.0018 (-0.56)	
<i>Dnews_tone_g</i>	0.0647*** (20.81)	
Controls	YES	YES
Stock /Time Fixed	YES	YES
N (million)	557.2257	557.2257
Under identification test		
SW Chi-sq	4022.86	
P-value	0.00	
Weak identification test		
Anderson-Rubin Wald F statistic	49.09	
Stock-Wright LM S statistic	840.20	
SW F statistic	446.98	

Note: This table presents the two-stage least squares (2SLS) regression results using the instrumental variable approach with $IV2_t$ constructed via the Lewbel (2012) heteroskedasticity-based method. num_t is the individual investors' information interaction intensity across two stock sub-forums on trading day t , CJ_t is the number of co-jumps on that day, *Controls* includes a set of control variables: Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and stock fixed effects. The instrument denoted with subscript “ $_g$ ” is generated via the Lewbel (2012) heteroskedasticity-based approach. This instrument does not appear in the structural equation and is used exclusively as an exclusion restriction to identify the

endogenous variable *num_p*. ***, **, and * denote the 1%, 5%, and 10% significance level, respectively.

Appendix E. Predictive Capacity Test of Co-jumps

We conduct a predictive capacity test of co-jumps on returns to determine the optimal holding time for positions in our trading strategies. It has been established that jumps contain substantial information, particularly in less mature financial markets. To examine whether co-jumps possess predictive power for future returns, we estimate the following regression models (D1) and (D2), with cumulative stock returns over the next one to three trading days as the dependent variables and upward and downward co-jumps as the independent variables.

$$R_{k,t+T} = a + b_1 up_CJ_{i,j,t} + \sum \gamma_{1l} Controls_{i,j,t} + \lambda_{i/j} + \tau_t + \varepsilon_{i/j,t} \quad (D1)$$

$$R_{k,t+T} = a + b_2 down_CJ_{i,j,t} + \sum \gamma_{2l} Controls_{i,j,t} + \lambda_{i/j} + \tau_t + \varepsilon_{i/j,t} \quad (D2)$$

where $R_{k,t+T}$ represents the future return of stock k , $T = 1, 2, 3$, and $k \in i, j$ for each pair (i, j) .

We divide the 2020–2023 dataset into eight consecutive six-month windows and conduct the predictive capacity test of co-jumps on returns for each time window separately. Tables D1–D8 present the predictive power of co-jumps across different windows. The empirical evidence reveals asymmetric predictive capabilities between upward and downward co-jumps: upward co-jumps significantly predict positive returns over the next one to two trading days in most windows, whereas downward co-jumps exhibit significant predictive power for negative returns on the following trading day in only a limited number of windows. Based on these findings, we determine the specific holding times for positions in the subsequent window.

Table E1: Predictive capacity of co-jumps (2020/01-2020/06)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0008*** (35.65)	0.0012*** (49.07)	0.0002*** (10.74)	0.0007*** (28.32)	0.0003*** (14.33)	0.0002*** (8.22)
R^2	0.2981	0.2979	0.2989	0.2981	0.2997	0.2985
N (million)	54.3715	54.3715	54.3715	54.3715	54.3715	54.3715
F	2369.6353	1803.2682	2192.9050	2152.3671	2155.3574	1927.4635
Panel B: Predictive capacity of downward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	0.0003*** (14.88)	-0.0002*** (-10.99)	-0.0001*** (-8.02)	0.0003*** (16.10)	0.0001*** (2.62)	-0.0005*** (-26.23)
R^2	0.2981	0.2979	0.2989	0.2981	0.2997	0.2985
N (million)	54.3715	54.3715	54.3715	54.3715	54.3715	54.3715
F	2223.6577	1575.1451	2179.8017	2117.6089	2135.0464	1974.8875
Note: Tables D1–D8 present the predictive power of co-jumps across different windows. R_i and R_j						

denote the returns of the stock pair (i,j) . Control variables include Dp , $Dvol$, $Dsize$, Dhs , $DInd$, $Dcity$, $same_holders$, $common_analysts$, and $Dnews_tone$. We also consider time and individual fixed effects. ***, **, and * denote the 1%, 5%, and 10% significance levels, respectively. The values in parentheses represent the t-statistics, computed based on standard errors clustered at the stock-pair level in all specifications, following Millo (2017). Similar notation and methodology apply to the results presented subsequently.

Table E2: Predictive capacity of co-jumps (2020/07-2020/12)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0013*** (38.01)	0.0008*** (26.87)	0.0001*** (3.16)	-0.0006*** (-20.42)	-0.0000 (-1.23)	-0.0000 (-0.12)
R^2	0.2444	0.2367	0.2441	0.2374	0.2445	0.2383
N (million)	58.3103	58.3103	58.3103	58.3103	58.3103	58.3103
F	2629.4117	1611.0137	2069.9922	1022.9298	1322.3536	817.2644
Panel B: Predictive capacity of downward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	-0.0005*** (-19.72)	-0.0001*** (-3.11)	-0.0002*** (-10.09)	0.0000 (1.08)	0.0006*** (26.03)	0.0003*** (13.93)
R^2	0.2444	0.2367	0.2441	0.2374	0.2445	0.2383
N (million)	58.3103	58.3103	58.3103	58.3103	58.3103	58.3103
F	2531.6253	1535.9713	2075.0919	981.3526	1383.9562	838.2556

Table E3: Predictive capacity of co-jumps (2021/01-2021/06)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0013*** (39.94)	0.0017*** (47.60)	0.0003*** (7.57)	0.0005*** (13.68)	-0.0001*** (-3.36)	0.0000 (0.98)
R^2	0.1192	0.1138	0.1201	0.1151	0.1207	0.1149
N (million)	64.4560	64.4560	64.4560	64.4560	64.4560	64.4560
F	4764.4735	2882.6486	3824.9951	2126.6317	3928.7596	1481.1930
Panel B: Predictive capacity of downward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	-0.0002*** (-8.74)	-0.0006*** (-22.03)	-0.0003*** (-10.18)	-0.0003*** (-10.64)	0.0000* (1.90)	0.0001*** (5.23)
R^2	0.1192	0.1138	0.1201	0.1151	0.1207	0.1149
N (million)	64.4560	64.4560	64.4560	64.4560	64.4560	64.4560

F	4614.7811	2700.3966	3828.7669	2111.0363	3930.2492	1486.2801
---	-----------	-----------	-----------	-----------	-----------	-----------

Table E4: Predictive capacity of co-jumps (2021/07-2021/12)

Panel A: Predictive capacity of upward co-jumps

	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0009*** (27.74)	0.0004*** (12.28)	0.0002*** (4.57)	-0.0002*** (-5.23)	-0.0004*** (-10.95)	-0.0002*** (-5.51)
R^2	0.1177	0.1127	0.1175	0.1127	0.1177	0.1129
N (million)	68.1264	68.1264	68.1264	68.1264	68.1264	68.1264
F	4316.0182	1379.9048	4494.5404	1969.1072	3873.7477	1452.1376

Panel B: Predictive capacity of downward co-jumps

	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	0.0002*** (6.21)	0.0001** (2.03)	0.0003*** (11.60)	0.0003*** (13.38)	0.0003*** (9.90)	0.0005*** (19.28)
R^2	0.1177	0.1127	0.1175	0.1127	0.1177	0.1129
N (million)	68.1264	68.1264	68.1264	68.1264	68.1264	68.1264
F	4264.2162	1361.5395	4506.1982	1983.7860	3869.5097	1493.7230

Table E5: Predictive capacity of co-jumps (2022/01-2022/06)

Panel A: Predictive capacity of upward co-jumps

	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0008*** (32.69)	0.0013*** (53.07)	0.0004*** (17.54)	0.0003*** (12.77)	-0.0001*** (-3.61)	0.0004*** (17.30)
R^2	0.2947	0.3012	0.2938	0.3007	0.2938	0.3006
N (million)	73.3340	73.3340	73.3340	73.3340	73.3340	73.3340
F	7351.7763	5694.6886	6888.7530	5267.4773	6017.6016	3873.8225

Panel B: Predictive capacity of downward co-jumps

	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	-0.0003*** (-12.41)	-0.0009*** (-42.58)	0.0001*** (4.60)	0.0004*** (19.31)	-0.0000 (-1.64)	-0.0002*** (-10.98)
R^2	0.2947	0.3012	0.2938	0.3007	0.2938	0.3006
N (million)	73.3340	73.3340	73.3340	73.3340	73.3340	73.3340
F	7271.5279	5584.8680	6865.1371	5284.3724	6017.5580	3855.4597

Table E6: Predictive capacity of co-jumps (2022/07-2022/12)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0006*** (29.48)	0.0008*** (37.13)	-0.0002*** (-8.70)	0.0002*** (9.45)	0.0002*** (9.78)	0.0003*** (12.26)
R^2	0.2119	0.2013	0.2131	0.2025	0.2151	0.2038
N (million)	77.2851	77.2851	77.2851	77.2851	77.2851	77.2851
F	5999.3698	4612.1468	4624.8027	3152.1412	4218.4915	2656.8512
Panel B: Predictive capacity of downward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	-0.0005*** (-23.56)	0.0003*** (16.07)	0.0007*** (35.06)	0.0009*** (42.39)	-0.0001*** (-4.85)	-0.0003*** (-13.49)
R^2	0.2119	0.2012	0.2131	0.2026	0.2151	0.2038
N (million)	77.2851	77.2851	77.2851	77.2851	77.2851	77.2851
F	5952.0900	4475.1800	4751.2671	3357.9616	4208.5182	2667.4272

Table E7: Predictive capacity of co-jumps (2023/01-2023/06)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0009*** (41.47)	0.0007*** (29.24)	0.0003*** (13.88)	0.0002*** (9.36)	-0.0003*** (-16.81)	-0.0001*** (-4.32)
R^2	0.1282	0.1312	0.1290	0.1325	0.1298	0.1335
N (million)	76.2067	76.2067	76.2067	76.2067	76.2067	76.2067
F	3903.7651	6453.7056	3837.3942	6105.1465	3254.8899	4903.3638
Panel B: Predictive capacity of downward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	-0.0002*** (-9.65)	-0.0002*** (-10.45)	-0.0002*** (-11.01)	-0.0001*** (-3.72)	0.0001*** (4.19)	0.0001*** (3.35)
R^2	0.162	0.166	0.165	0.165	0.163	0.167
N (million)	76.2067	76.2067	76.2067	76.2067	76.2067	76.2067
F	3746.0272	6398.9102	3831.1719	6092.1992	3229.8935	4900.4222

Table E8: Predictive capacity of co-jumps (2023/07-2023/12)

Panel A: Predictive capacity of upward co-jumps						
	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
up_CJ_t	0.0003***	0.0005***	-0.0004***	-0.0001***	0.0000***	0.0000

	(14.56)	(23.96)	(-25.09)	(-7.18)	(2.69)	(1.32)
R ²	0.1757	0.1768	0.1773	0.1775	0.1797	0.1791
N (million)	78.2105	78.2105	78.2105	78.2105	78.2105	78.2105
F	7235.0607	9571.4874	5421.6501	7910.3480	2898.0976	5368.5991

Panel B: Predictive capacity of downward co-jumps

	t+1		t+2		t+3	
	R_i	R_j	R_i	R_j	R_i	R_j
$down_CJ_t$	0.0001***	0.0000***	0.0006***	0.0006***	0.0003***	0.0006***
	(10.04)	(2.84)	(42.65)	(37.65)	(19.98)	(38.24)
R ²	0.1757	0.1768	0.1773	0.1775	0.1797	0.1791
N (million)	78.2105	78.2105	78.2105	78.2105	78.2105	78.2105
F	7221.4409	9533.9413	5545.2672	8035.9018	2930.5155	5502.3438

Appendix F. Implementation Details of the Strategy

The implementation details of the strategy are as follows:

(i) **Stock Pair Classification:** In each window, based on the results of market capitalization heterogeneity analysis, we classify stock pairs into two categories: “Big-Big” if both stocks in the pair are “Big-Cap” stocks, and “Small-Small” if both stocks are “Small-Cap” stocks on every trading day.

(ii) **Daily Trading Decisions:** Within each category, we first rank the stock pairs based on their positive or negative individual investors’ information interaction intensities. Long positions are then established in the top 100 stock pairs with the highest positive information interaction intensity on each trading day, with all stocks equally weighted. Similarly, short positions are taken in the top 100 stock pairs exhibiting the highest negative information interaction intensity on each trading day, also under an equal-weighting scheme. These positions are closed at the end of the holding time.

(iii) **Holding Time Determination:** The initial holding time for each position is set to one trading day. To maximize profitability, we adjust the holding times for positions in each window based on the predictive capacity test results of co-jumps on future returns from the preceding window. For example, the predictive capacity test results for the January-June 2023 window show that upward (downward) co-jumps have significant predictive power for positive (negative) returns over the subsequent two trading days. Accordingly, for the July-December 2023 window, the holding time for each position is determined as follows: if an upward co-jump is detected in a stock held in a long position, the holding time for that stock is extended to two trading days; similarly, a downward co-jump in a stock held in a short position extends the holding time to two trading days for that stock. Given that the predictive capacity of co-jumps varies across time windows, the holding times for both long and short positions are dynamically adjusted according to their predictive accuracy in each specific window, as shown in Table F.

Table F. The Holding Time for Long and Short Positions

	2020.07- 2020.12	2021.01- 2021.06	2021.07- 2021.12	2022.01- 2022.06	2022.07- 2022.12	2023.01- 2023.06	2023.07- 2023.12
Holding time_L	3 days	1 day	2 days	1 day	2 days	1 day	2 days
Holding time_S	0 day	1 day	2 days	0 day	1 day	0 day	2 days

Note: This table presents the optimal holding times for long and short trading positions across different windows. Holding time_L represents the extended holding days for long positions, while Holding time_S represents the extended holding days for short positions, as derived from the regression results of upward and downward co-jumps on future returns.