Return Predictability and Online Stock Opinions Published During Trading and Non-Trading Hours

Jia Wu, Junjie Ren, and Lin Huang*

ABSTRACT

This study analyzes the impact of trading- and non-trading-hour opinions on returns using data collected from an online stock forum in China. We find that non-trading-hour opinions have a stronger influence on returns than trading-hour opinions. However, a return reversal is observed during the subsequent trading periods based on non-trading-hour opinions, suggesting a tug-of-war between individual investors and arbitrageurs. Additionally, the effect of non-trading-hour opinions on returns is higher when firms announce important events overnight. These opinions also attract more investor attention. We propose that the announcement of such events exposes investors to high levels of uncertainty, leading them to seek advice through online forums. Our analysis suggests that investor sentiment and value-relevant information contained in online articles are likely factors that contribute to the return predictability of these opinions (JEL codes G12, G14).

Keywords: Online stock forum, return predictability, trading hours, non-trading hours

* Jia Wu, Department of Economics, Shandong University; Weihai. Institute for Interdisciplinary Research, Weihai, China. Junjie Ren, Department of Economics, National University of Singapore, Singapore. Lin Huang, the corresponding author, Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu, China (Email: lhuang@swufe.edu.cn).
The authors thank many colleagues and friends for their helpful comments. We acknowledge the funding support provided by the National Natural Science Foundation of China (Grant No. 72073051), and the Natural Science Foundation of Guangdong Province (Grant Nos. 2020A1515010421 and 2021A1515012304).

1. Introduction

A growing body of literature finds that online opinions can forecast future stock returns (Chen et al., 2014; Huang et al., 2016; Jiang et al., 2018).[1] This means that investors may rely on others' opinions shared through online social media to make decisions about buying or selling stocks, but how do they respond to such opinions issued during trading and non-trading hours? An important fact is that companies often disclose a substantial amount of material public information after the market closes to avoid fluctuation in stock prices (Barclay and Hendershott, 2003). Do opinions published overnight have a greater ability to predict returns than those published during trading hours? Moreover, empirical evidence from previous studies has shown that overnight returns perform differently from intraday returns due to the "tug-of-war" between individual investors and arbitrageurs (Akbas et al., 2022; Lou et al., 2019, 2022). Since online platforms serve as ideal places for individual investors to seek advice and share information,[2] comparing the effects of trading and non-trading-hour opinions on returns will provide valuable insights into return patterns. Answering this question can also shed light on how investors access information during these periods.

This study fills a gap in the literature by using a dataset collected in the most popular online stock discussion forums in China (i.e., *Guba Eastmoney*). Our data enable us to identify the number of articles and the related number of comments and reads published during trading or non-trading hours on this forum so that we can assess the return predictability of opinions published during trading- and non-trading hours. China provides an ideal setting for this research because, as the largest emerging capital market in the world, the Chinese stock market incorporates the majority of retail investors.[3] Since retail investors probably lack the knowledge to properly interpret such information, they tend to seek the opinions of others on online forums to guide their investments. This study also

[1] Some studies have found different results. For example, Tumarkin and Whitelaw (2001), Antweiler and Frank (2004), and Das and Chen (2007) find no evidence of the relationship between online reviews and stock returns. Also, Kim and Kim (2014) find no evidence that the investor sentiments extracted from internet postings can forecast future stock returns, volatilities, or trading volumes.

[2] A Cogent Research report published in 2013 shows that approximately one-third (34%) of affluent investors in the United States (US) use social media platforms for personal finance and investing purposes (see www.businesswire.com/news/home/20130222005037/en/Cogent-Research-Trending---Social-Media-Fuels). As early as 2012, the US Securities and Exchange Commission alerted investors to the risks associated with using social media, which they described as "landscape-shifting" (SEC, 2012).

3 An investor survey conducted by the Shenzhen Stock Exchange shows that as of 2018, small and medium investors with assets invested in stocks less than RMB 500,000 (approximately 72,000 USD) account for 80% of the overall individual investors in the market (see www.szse.cn/English/about/news/szse/t20190319_565573.html).

focuses on opinions in different tones.[4] Our analysis is based on nearly three million observations related to 3,154 listed firms between 2008 and 2021, which grant us sufficient statistical power to detect the potential effect.

Our findings can be summarized as follows. The proportion of positive and pessimistic opinions can significantly predict abnormal returns over the next few days. Non-trading-hour pessimistic opinions can only affect stock returns on the next day, while trading-hour and non-trading-hour positive opinions have persistent effects. Our long-term analysis of cumulative abnormal returns also supports these results. We sharpen our analysis by decomposing daily returns (close-to-close) into overnight returns (close-to-open) and daytime returns (open-to-close). We find that *Guba* opinions primarily impact overnight returns. Non-trading-hour opinions have a more pronounced effect than trading-hour opinions. However, a substantial portion of the returns affected by non-trading-hour opinions reverses during the next trading hours. This result can be explained by the fact that individual investors tend to trade near the opening of the market, while institutional investors are likely to trade in the opposite direction during the trading day (Berkman et al., 2012; Akbas et al., 2022; Lou et al., 2019, 2022). Given that individual investors are the main users of *Guba* forum, our results are consistent with the literature showing that individual and institutional investors represent two distinct clientele and trade in opposite directions.

Second, we find that investor sentiment and value-relevant information contained in *Guba* articles are likely channels of the predictability of *Guba* opinions. We conduct two exercises. First, we explore whether *Guba* opinions act as a proxy for investor sentiment by comparing the predictive power of opinions on stocks with different market values. Our logic follows Baker and Wurgler (2006), who show that sentiment has a stronger impact on stocks that are difficult to evaluate, such as small stocks. However, our findings indicate that non-trading-hour opinions lead to a higher return for big stocks, suggesting that investor sentiment seems not to be the sole driving force behind the return predictability of opinions. Second, we follow Chen et al. (2014) to examine whether *Guba* opinions can predict earnings surprises. We discover that *Guba* opinions are indeed able to predict

4 Literature shows that investors may respond differently to positive and negative news. For example, Epstein and Schneider (2003, 2008) and Williams (2009) find that investors are more sensitive to negative news.

earnings surprises, but this effect only exists for non-trading-hour opinions and among big stocks. Since no value-relevant information was released about small firms, this suggests that the effects of *Guba* opinions for small firms are primarily driven by the sentiment of noise investors.

We continue our analysis by exploring possible channels to understand the different effects between trading- and non-trading-hour opinions. Literature shows that firms tend to disclose substantial material information after the market closes (Barclay and Hendershott, 2003; Santosh, 2016). In our sample period, we find that approximately 85% of important information, known as "major events," is published by listed firms after the market closes. We first show that opinions can still affect returns even if we control for the effects associated with major-event days, implying that the predictive power of *Guba* opinions cannot be fully explained by these major events or the information they contain. We then compare the return predictability between opinions published on major-event days and those published on regular days. Our findings reveal that the predictability of *Guba* opinions on stock returns is significantly larger for opinions published on days when firms announce major events. Finally, we show that opinions published on major-event days attract more investor attention, as measured by the number of articles published and the number of comments made on the forum. We propose a plausible explanation, that the announcement of these events leads investors to a higher level of uncertainty, which causes them to seek the opinions from others on the online forums, resulting in the return predictability of those opinions.

We further discuss two related issues. First, we investigate whether *Guba* opinions can predict returns under different market conditions. We find that positive opinions tend to have higher effects on stock returns during good times, implying that investors tend to overestimate positive opinions when the stock market surges. Second, we test whether investors have an opportunity to arbitrage by adopting a strategy that leverages the opinions published on the *Guba* forum. After incorporating transaction costs into portfolio returns, we find that investors can barely make profits by using the "opinion strategy."

This study relates to the literature on the stock market and the informational role of large crowds (e.g., Da et al., 2011; Kelley and Tetlock, 2013; Chen et al., 2014; Huang, 2018). Among these studies, the closest work to ours is Chen et al. (2014), who find that the views expressed in articles and commentaries on online forums can predict future stock returns. The authors indicate that social media platforms allow investors to directly and immediately interact and exchange information with one another, and such information exchange can become an important source of information for stock pricing and potentially create a “wisdom of crowds.” Like Chen et al. (2014), we find that *Guba* articles contain value-relevant information, but this information only concerns big firms. We add value to Chen et al.’s (2014) study by distinguishing the opinions published during trading and non-trading hours. We also reveal that retail investors who encounter the release of important firm-related news, which is usually announced after the market closes, are more likely refer to the online stock forum for advice, helping us to understand the role of the online forum in price fluctuation. Our work relates to a growing body of literature on social forums and their influence on stock prices (Antweiler and Frank, 2004; Leung and Ton, 2015; Renault, 2017; Hirshleifer, 2020). Our study complements these studies by providing new evidence from the largest emerging market in the world to highlight the predictive power of peer opinions published during trading and non-trading hours.

We build on the recent literature showing that individual investors and institutional investors are different types of investors. Individual investors are inclined to trade overnight and close to the opening of the market, while institutional investors trade during the daytime in an opposite direction, thereby offsetting the return changes caused by the individual investors (Akbas et al., 2022; Berkman et al., 2012). We contribute to the existing literature by showing that non-trading-hour articles published on the *Guba* forum, which is favored by individual investors, have a significant impact on overnight returns. Additionally, we find evidence of noticeable reversals in daytime returns, suggesting the influence of institutional investors during trading periods. These findings have important implications regarding the role of online forums as important sources of information for individual investors, particularly when they are faced with uncertainty.

This study contributes to the existing literature on investor sentiment (for an overview, see Baker and Wurgler, 2007; Baker and Wurgler, 2006; García, 2013). The main focus of this literature is the role of investor sentiment on stock returns. For example, García (2013) collects words from two financial columns from the *New York Times* and shows that these words mainly serve as a proxy for market sentiment. Huang et al. (2016) and Jiang et al.

(2018) find similar results, showing that online posts mobilize crowds and induce herd behaviors in terms of investment actions. On the contrary, Chen et al. (2014) find that social media platforms contain value-relevant information that has not been incorporated into stock prices. We advance this strand of the literature by showing that the power of *Guba* articles in forecasting stock returns can be attributed to both investor sentiment and the information embedded within these articles.

Finally, this study is also related to a body of research that investigates the effect of the *Guba Eastmoney* forum on the Chinese stock market. *Guba Eastmoney*, as a representative online community of Chinese retail investors, has attracted great attention recently. For example, Huang et al. (2016) examine the messages posted on *Guba Eastmoney* and find that individual investors pay more attention to stocks of local companies than those of non-local companies. You et al. (2017) use this forum to study the corporate governance roles of state-controlled and market-oriented media separately from social media. Jiang et al. (2018) find evidence of investor communication in *Guba Eastmoney* and the co-movement of related stocks. What sets our study apart from the above literature is that we show new evidence that investors rely on online forums but act differently in response to trading- and non-trading-hour articles.

The remainder of this paper is structured as follows. Section 2 describes our data. Section 3 presents the model specification and the main estimation results. Section 4 explains our results. Section 5 discusses related issues and provides robustness tests. Section 6 concludes the paper.

2. Data

We obtain *Guba* data from the Chinese Research Data Services (CNRDS).[5] We also collect company news published in Chinese newspapers and stock analyst data from CNRDS. We retrieve stock returns and corporate financial information from the RESSET Financial Research Database (RESSET/DB).[6]

5 See http://www.cnrds.com.
6 See http://www.resset.cn/endatabases.

2.1 *Guba* opinions

Eastmoney.com is the most popular investment-related website in China. The *Guba* forum is an online discussion community on *Eastmoney.com*, that allows users to express their opinions and communicate with others regarding stock investment.[7] *Guba* can be viewed as a Chinese version of Seeking Alpha and is the most popular stock-related investment forum in China.[8] According to website traffic statistics from SimilarWeb and Alexa, *Eastmoney.com* has accumulated 54.35 million visits per month as of February 2018, making it the 6th largest investment website in the world. The *Guba* forum significantly contributes to this traffic, with approximately 49% of *Eastmoney.com* visitors accessing this website. With such a high volume of traffic, we suppose that *Guba* is widely used by retail investors.

Each stock has a separate subforum in *Guba* and is indexed uniquely by the name of a firm along with its six-digit stock code. Investors need to enter the subforum for a specific stock to browse related articles. Registration in *Guba* is free of charge. Reading articles on *Guba* does not require registration, but only registered users can publish new articles and comment on existing ones. Figures A1 and A2 in the Appendix present a list of sample articles for a specific stock and comments placed under an article. Articles published by institutional investors and the listed company tend to attract more reads and comments. As shown in Figure A1, the 8th article was published by an institutional investor, while the 11th and the 12th articles were published by *Wanke A-stock News*,[9] the official account of the company that owns this stock. These articles receive a higher number of reads and comments compared with the others. Moreover, after browsing more pages, we observe that the company's official user, the *Wanke A-stock News*, often releases its latest operation news outside trading hours, because it publishes such news on *Guba* forum simultaneously with its companies' own announcements, which are usually made after the market closes.

In the A-share market of China, the auction bidding from 9:15 AM to 9:30 AM determines the opening prices of stocks. Trading starts at 9:30 AM, stops between 11:30 AM and 1:00 PM, and closes at 3:00 PM.[10] We treat the trading hours as continuous hours

7 *Guba* means stock forum when translated into Chinese.
8 See http://www. seekingalpha.com.
9 By looking for the "v" identifier next to a poster's ID, we can generally determine if it is a retail investor, institutional investor, or a listed company. This is because verified investors are obligated to disclose their institution's name.
10 The auction bidding from 9:15 AM to 9:30 AM determines the opening prices of stocks.

for each day. Specifically, we define trading-hour opinions as those posted from 9:00 AM to 3:00 PM and non-trading-hour opinions as those posted from 3:00 PM to 9:00 AM on the next trading day (i.e., between the market closing on a trading day and the market opening on the next trading day).[11] Table A1 in the Appendix shows the trading and non-trading hours.

The CNRDS collects all articles, their comments, and reading times published on *Guba* forum from 2008. It then constructs a dictionary and utilizes a machine-learning process to categorize these articles and comments into positive, neutral, and pessimistic ones (see Appendix for details). The *Guba* forum records the precise time when all articles and their corresponding comments are made. Using this information, we can determine the number of positive and pessimistic articles published during trading and non-trading hours on each day.

Our main explanatory variables are constructed based on this information, namely, proportions of pessimistic and positive opinions on stock *i* to the total number of trading-hour or non-trading-hour opinions on this stock on day *t*. Specifically, for non-trading-hour opinions, we define:

$$NPess_{it} = \frac{\sum_k NPessOpinion_{ikt}}{\sum_k NOpinion_{ikt}} \quad \text{and} \quad NPos_{it} = \frac{\sum_k NPosOpinion_{ikt}}{\sum_k NOpinion_{ikt}},$$

For trading-hour opinions, we define:

$$TPess_{it} = \frac{\sum_k TPessOpinion_{ikt}}{\sum_k TOpinion_{ikt}} \quad \text{and} \quad TPos_{it} = \frac{\sum_k TPosOpinion_{ikt}}{\sum_k TOpinion_{ikt}},$$ [12]

where *i* represents the stock and *t* denotes the day. Every day, stock *i* receives several opinions, and we use *k* to denote each opinion. $NPessOpinion_{it}$ ($TPessOpinion_{it}$) is the total number of non-trading-hour (trading-hour) pessimistic opinions about stock *i* on day *t*. $NPosOpinion_{it}$ ($TPosOpinion_{it}$) is the total number of non-trading-hour (trading-hour) positive opinions about stock i published on day *t*. $NOpinion_{it}$ ($TOpinion_{it}$) denotes the total number of non-trading-hour (trading-hour) opinions about stock *i* published on day *t*. We refer to $NPess_{it}$, $NPos_{it}$, $TPess_{it}$, and $TPos_{it}$ as *Guba* opinion

11 Articles published on weekends and holidays are treated similarly. We define articles published after the market closes on the last trading day before the weekends or holidays and before the market opens on the next trading day as posted during non -trading hours of the last trading day.

12 The purpose of this paper is to examine whether opinions published on day *t* can forecast future returns starting from day *t*+1. We construct our four opinion variables based on the information that is available on day *t*. That means our strategy avoids the issue of look-ahead bias.

variables. Opinions that are not positive and pessimistic are grouped into neutral ones. We treat the proportion of neutral opinions as a reference group.

Our original dataset includes 3,154 non-financial stocks traded over a period of 3,407 days.[13] We have excluded stocks on a particular day that received no opinions (nearly 1,600,000 observations are dropped). To account for the effect of abnormal returns in the previous days, we control for lagged abnormal returns two days before the opinions are published up to the day they are published, and we exclude approximately 1,100,000 observations as a result. After applying these data screening methods, we are left with a sample of 3,422,599 stock-by-date observations. In some specifications, we also control for the number of newspaper articles. This gives us a smaller sample of 2,974,797 observations.

Wo obtain the number of reading times and comments of articles that a stock received each day from the CNRDS. We use the above two variables to measure investors' attention. We match the opinions, reading times, and comments with stock returns according to the company names and trading dates.

Table 1 presents the summary statistics for the aforementioned variables. For pessimistic opinions, nearly 28.9% are published during trading hours and 18.4% are published during non-trading hours. For positive opinions, nearly 29% are published during trading hours and 31.8% are published during non-trading hours. The standard deviations of the proportions of positive and pessimistic opinions slightly increase for those published during non-trading hours, indicating a moderate increase in the divergence of investor attitudes after the market closes.

On average, 29.8 and 16.7 articles are published for a stock during trading-hour and non-trading hours, respectively. However, a non-trading-hour article receives more reads (1,400 vs. 1,000; p-value=0.000) and comments (2.43 vs. 1.95; p-value=0.000) than trading-hour opinions on average, probably due to the fact that more institutional investors and listed companies publish articles during non-trading hours. We will show that firms usually announce major events about their operations after the market closes and that these events tend to raise investor interest and discussions on *Guba*.

13 On average, a stock is traded for 1,946 days in our sample when considering the delisted and newly listed stocks in later years. We do not exclude those firms with extremely low-priced stocks. Our results are quantitively unchanged when these companies are excluded.

2.2 Abnormal returns

Our main dependent variables are abnormal returns in the next 5 trading days (short-run outcomes) and cumulative abnormal returns (CARs) over the 5-, 10-, 20-, 60-, and 120-day windows (long-run outcomes) after the opinions are published. Daily returns refer to the price change from the market close on day *t* to the market close on day *t*+1 (i.e., close-to-close return). The abnormal returns are computed as the difference between the raw return of a stock and the return explained by the five-factor model of Fama and French (2015). The CARs are calculated by summarizing the abnormal returns in corresponding windows.[14] For most of our regressions, we control for abnormal returns of stock *i* from 2 days before the opinions are published to the day when opinions are published. We also control for the past three-month holding period returns of stock *i*. Panel A of Table 2 reports the descriptive statistics of these returns.

2.3 Other variables

To disentangle the effects of *Guba* opinions from those of news released through traditional media, we collect company news published by financial and economic newspapers in China. News articles are collected from eight nationwide financial newspapers in China, namely, *China Securities Journal*, *Shanghai Securities News*, *First Financial Daily*, *21st Century Business Herald*, *China Business Journal*, *The Economic Observer*, *Securities Daily*, and *Securities Times*. These newspapers are either owned or controlled by the state, financial institutions, public companies, or wealthy individuals and are renowned in China for their timely reporting and high-quality news. Therefore, these newspapers play a key role for investors to obtain financial information (You, Zhang, and Zhang, 2017). Similar to *Guba* opinion data, CNRDS categorizes newspaper articles as neutral, positive, and pessimistic. In most of our results, we control for a dummy variable indicating whether a firm is covered by the aforementioned newspapers, the total number of newspaper articles, as well as the number of pessimistic and positive newspaper articles.

We will show that firms' announcement of major events after market close is the main reason that incurs the different effects on returns between trading- and non-trading-hour

14. CAR would not be set to a missing value if AR is missing for some days within the window that used to calculate CAR. This explains why we observe different sample sizes when using AR and CAR as dependent variables even when focusing on the same period (e.g., $Ret_{i,t+5}$ and $CAR_{i,t+5}$). Our results are not sensitive to this calculation method for CARs.

opinions. We identify the hours and dates when companies hold conferences to announce major events. Major events are defined as those events that are expected to greatly affect share prices, such as the appointment of a new CEO, change in dividend policies, acquisition, etc.[15] To control for compound effects of other events, we identify the date when companies publish financial reports. We also calculate earnings surprise to test whether the opinions published on *Guba* include value-relevant information. We use the difference between the reported EPS and the average EPS forecasted by analysts and the difference between the reported profits and the average profits forecasted by analysts to calculate earnings surprise. We also collect information on firm size.

Panel B of Table 2 presents the summary statistics for the firms' characteristics and newspaper variables. The average market value, which is measured in logarithms, is 22.62. The average book-to-market ratio stands at 0.46. On average, these firms generate annual revenues and profits that account for 43.9% and 1.87% of their market values, respectively. Firms invest 3.4% of their revenues in R&D expenditures on average. Throughout our sample period, the average daily trading volume amounts to 213.1 million RMB. Approximately 4.6% of the listed firms are covered by the newspapers. A firm is reported by 0.07 newspaper articles on a day, of which 0.03 is positive articles and 0.03 is pessimistic articles.[16] A firm is averagely covered by 0.23 analysts. On average, a firm discloses 0.09 major events on a single day. Regarding the earnings surprise variables, they have average values of −0.001 and −0.06, depending on the different measurements of earnings.

3. Model estimation and main results

3.1 Model estimation

We estimate the effect of opinions published on *Guba* on abnormal returns over the subsequent trading days and the long-run CARs using the following model:

$$Ret_{i,t+\tau} = \alpha + \beta_1 NPess_{it} + \beta_2 NPos_{it} + \beta_3 TPess_{it} + \beta_4 TPos_{it}$$
$$+\mathbf{X}_{it}\boldsymbol{\delta} + \mu_i + \mu_{my} + \varepsilon_{it}, \qquad (1)$$

15 Table A2 in the online Appendix lists these major events.
16 The number of neutral news is excluded.

where $Ret_{i,t+\tau}$ is the abnormal returns or the CARs for stock *i*. In our short-run analysis, we analyze the abnormal returns τ days after the *Guba* opinions are published. In our long-run analysis, we analyze the CARs, which are calculated over the τ window after the *Guba* opinions are published. $NPess_{it}$ and $NPos_{it}$ are the proportions of pessimistic and positive articles on firm *i* posted during non-trading hours on day *t*, and $TPess_{it}$ and $TPos_{it}$ are the proportions of pessimistic and positive opinions on firm *i* posted in the trading hours on day *t*. $\mathbf{X}_{it}$ denotes a set of control variables, including the abnormal returns two days before the opinions are published to the day they are published ($ARet_{it}$, $ARet_{i,t-1}$, $ARet_{i,t-2}$), the three-month holding period abnormal returns ($CAR_{i,t-60,t-3}$), and the number of reading times and comments. We also control a dummy indicating whether a firm is mentioned by newspapers on day t ($INews_{it}$), the total number of newspaper articles in which firm *i* is mentioned on day *t* ($Newsnum_{it}$), and the number of pessimistic and positive newspaper articles about firm *i* on day *t* ($NegNews_{it}$ and $PosNews_{it}$). μ_i is a set of dummies indicating for firm fixed effects. μ_{my} is a set of dummies indicating year-by-month fixed effects, which are used to control for any heterogeneity due to month and year. ε_{it} is the error term.

The estimated coefficients β_1 to β_4 measure the effects of trading- and non-trading-hour *Guba* opinions on stock returns in the future. Notably, those estimated coefficients are exempt from reverse causality because the proportion of positive and negative opinions are measured before the stock returns. To deal with the possible serial correlation over time, we cluster standard errors at the firm level.

3.2 Main results

In this section, we initially conduct the short-run analysis by examining the abnormal returns published a few days after the opinions are published. Afterward, we conduct the long-run analysis by examining the CARs from 5- to 120-day windows after the opinions are published. To enhance the readability of the results table, we multiply the estimated coefficients by 1,000 when abnormal returns or CARs are used as the dependent variables.[17]

17 Such adjustment is applied to results in Tables 3–5, 7, 10, A3–A5, and A8.

A. Abnormal returns in a week (short-run evidence)

The estimated results for abnormal returns on the next trading day are reported in columns 1–5 of Table 3. Column 1 shows the effects of non-trading-hour opinions on the next-day stock returns. The estimated coefficient of $NPess_{it}$ is −0.752 ($p < 0.01$), implying that the abnormal returns decrease by 0.02% (0.02% = 0.075% × [0.25−0]) when the proportion of non-trading-hour pessimistic opinions increases from the 25th to the 75th percentile.[18] Non-trading-hour positive opinions significantly predict next-day returns as well. Specifically, the abnormal returns increase by 0.015% (0.015% = 0.047%× [0.5−0.19]) when the proportion of positive opinions increases from the 25th to the 75th percentile.[19] The *p*-value of the *t* statistic for testing the equality of magnitudes of the effects between positive and negative opinions is close to 0, indicating that the non-trading-hour pessimistic and positive opinions statistically differ from each other in terms of predictability on next-day return. Column 2 shows the effects of the trading-hour opinions on next-day returns. The abnormal returns are 0.01% lower when the proportion of pessimistic opinions increases from the 25th to the 75th percentile and are 0.01% higher if the proportion of positive opinions rises from the 25th to the 75th percentile.[20] When testing the equality of magnitudes of the coefficients between positive and negative opinions, we obtain a large *p*-value, implying that we cannot reject the null hypothesis that the two coefficients have the same absolute value.

In column 3, we perform a horse race by simultaneously estimating the trading- and non-trading-hour opinions in a regression. The estimated coefficients of the *Guba* opinion variables are quite similar to those shown in the previous columns. In this specification, we can test whether trading- and non-trading-hour opinions have different predictive powers for stock returns. The small *p*-values of the *t*-statistic for testing the equality of the coefficients of $NPess_{it}$ and $TPess_{it}$ (p=0.007) and the equality of the absolute value of the coefficients of $NPos_{it}$ and $NPos_{it}$ (p=0) indicate that the non-trading-hour opinions have a higher influence on next-day returns compared with the trading-hour opinions. One

18 The 25th and 75th percentiles of the proportion of pessimistic articles published during non-trading hours are 0 and 0.25, respectively.
19 The 25th and 75th percentiles of the proportion of positive articles published during non-trading hours are 0.19 and 0.5, respectively.
20 The 25th and 75th percentiles of the proportion of pessimistic articles published during trading hours are 0.17 and 0.36. The 25th and 75th percentiles of the proportion of positive articles published during trading hours are 0.17 and 0.35.

possible reason is that *Guba* forum is more favored by retail investors, who are more sensitive to those online opinions.

One might concern that the return predictability of *Guba* opinions arises from the compound effects of newspaper articles. For example, investors may forward the information in newspaper articles to *Guba* or discuss stocks inspired by what they have read in newspapers. In column 4, we further control whether a stock is reported by newspapers on the same day as the opinions are published on *Guba*. Being reported by a newspaper increases the abnormal returns by 0.012% (t=2.84) on the following day. This positive effect may be due to investors paying more attention to those stocks that are reported in newspapers, which leads to a net increase in the purchase of these stocks (Barber and Odean, 2008; Da et al., 2011).[21] The estimated coefficients on opinion variables are consistent with those in column 3, although the coefficients on non-trading-hour opinions are slightly smaller. In column 5, we further control the number of positive and negative newspaper articles and the total number of newspaper articles about a stock. A greater number of positive newspaper articles is linked to a higher return on the following day, whereas an increase in negative and neutral newspaper articles leads to a decrease in returns on the subsequent day. After controlling for these newspaper variables, we still find that the estimated coefficients on our opinion variables are significant, suggesting that the return predictability of *Guba* opinions is not driven by the information published in newspapers.

In the above analysis, we use daily return (close-to-open return) as our dependent variable. To provide insights into the source of next-day return changes related to *Guba* opinions, we follow the method suggested by Akbas et al. (2022) to decompose daily returns into overnight (close-to-open) and daytime (open-to-close) returns. Afterward, we calculate the corresponding abnormal returns using the five-factor model of Fama and French (2015). Table 4 presents the results of using overnight and daytime returns as dependent variables. In this table, the odd columns control for abnormal returns from two days before the opinions are published to the day they are published, the last three-month

21 Our explanation aligns with the attention theory of Barber and Odean's (2008), which suggests that individual investors tend to be net buyers of stocks that attract their attention. In our context, investors may pay more attention to those stocks that have been covered by newspapers, resulting a high increase in return on the following day. This theory is supported by later empirical works, such as Da et al. (2011).

holding period abnormal returns, the number of reading times and comments, and the firm and year-by-month fixed effects. In the even columns, we further control for a dummy indicating that firm *i* is covered by newspapers, the number of firm-*i* related pessimistic, and positive articles, and the total number of newspaper articles published on day *t*. To facilitate comparison, columns 1–2 present the results from columns 3 and 5 in Table 3. We conduct three exercises.

First, in columns 3–4, we regress overnight returns on day *t* on trading-hour positive and pessimistic opinions on day *t*. Results reveal a significant decrease (increase) in overnight returns when the proportion of trading-hour pessimistic (positive) opinions increases. Second, in columns 5–6, we regress overnight returns on day *t* on trading-hour opinions on day *t* and the contemporary overnight non-trading-hour opinions on day *t*. We find that non-trading-hour opinions can significantly predict overnight returns, and their effects on overnight returns are notably larger than those of trading-hour opinions. Specifically, when we control for newspaper variables, column 6 shows that the overnight returns increase by 0.14% when the proportion of non-trading-hour pessimistic opinions increases from 0 to 100%, but only increase by 0.04% when the proportion of pessimistic opinions increases from 0 to 100%. This result suggests that non-trading-hour opinions have a greater impact on overnight returns than trading-hour opinions.

Third, in columns 7–8, we regress daytime returns on day *t*+1 on trading-hour and non-trading-hour opinions on day *t*. We notice that the returns moved by non-trading-hour opinions reverse on the next trading day. Specifically, columns 7–8 show that nearly one-half (one-fifth) of the returns reverse following the release of non-trading-hour pessimistic (positive) opinions. Previous trading-hour opinions only have a small effect on daytime returns. These findings are in line with the literature showing that daytime returns negatively correlate to previous overnight returns (Akbas et al., 2022). A popular explanation is that individual investors are inclined to trade overnight, while institutional investors tend to trade during trading days to engage in arbitrage activities against the overnight price fluctuations caused by individual investors (Berkman et al., 2012). As alluded to above, the *Guba* forum is widely used by Chinese retail investors. The literature also shows that individual and institutional investors represent two investor clientele that cause opposing price pressures during their respective overnight and trading day periods

(Akbas et al., 2022; Lou et al., 2019, 2022). By comparison, no daytime reversals are observed following the release of trading-hour opinions, which also supports the hypothesis that daytime reversals are against the overnight price fluctuations caused by individual investors who tend to follow non-trading-hour opinions to trade.

We then determine whether *Guba* opinions can predict the stock returns beyond the next trading day. Columns 6–9 in Table 3 show the estimated effects of *Guba* opinions on the abnormal returns from days *t*+2 to *t*+5. The coefficients on $NPess_{it}$ are small and insignificant from days *t*+2 to *t*+5, implying that non-trading-hour pessimistic opinions can only predict next-day returns. In comparison, the estimated coefficients on $NPos_{it}$, $TPess_{it}$, and $TPos_{it}$ are significant and maintain the same sign across columns 5–9, suggesting that trading-hour opinions and non-trading-hour positive opinions have a persistent effect on abnormal returns at least in the subsequent five trading days.

Following the Reviewer's advice, we conduct additional analysis to compare the effects of pessimistic opinions with other ones. First, we only include the proportion of pessimistic opinions in Equation (1). Our results are consistent, except that the estimated coefficients on pessimistic opinions are close to the differences between the corresponding coefficients in Table 3, namely, between the coefficients of pessimistic and positive opinions (Panel A of Table A8 in the Appendix). Second, following García's (2013) approach, we construct pessimism factors by subtracting the proportion of positive opinions from the proportion of pessimistic opinions. We then replace our opinion variables with these pessimism factors based on Equation (1). The estimated coefficients on these factors represent the impact of a 100% increase in net pessimistic opinions on stock returns. Our findings are in line with those presented in Table 3 (Panel B of Table A8 in the Appendix).

To put the results in Table 3 into perspective, Figure 1 illustrates the dynamics of the returns as predicted by *Guba* opinions from days *t*+1 to *t*+5, including the estimated coefficients and 95% confidence intervals. The top left figure presents a return reversal that starts on the 5th day after the non-trading-hour negative *Guba* opinions have been published. The top right figure shows that the returns are persistently negative from days *t*+1 to *t*+5, suggesting a stable predictive power of the trading-hour pessimistic opinions. The bottom two figures show that the trading- and non-trading-hour positive opinions can only predict positive returns on day *t*+1.

B. CARs in half a year (long-run evidence)

To estimate the return predictability of *Guba* opinions in longer periods, we examine the CARs in windows over the 5, 10, 20, 60, and 120 trading days after the *Guba* opinions are published. We then regress CARs on the proportion of positive/pessimistic *Guba* opinions published during trading and non-trading hours. The same covariates in column 9 of Table 3 are controlled. Figure 2 presents the estimated coefficients and their 95% confidence intervals (Table A3 in the Appendix shows the estimated coefficients). Noted that the interpretation of the estimated coefficients for CARs is different from those for the abnormal returns. For example, if the returns exhibit reversals after the posting of articles, then the estimated coefficients for CARs will move to 0 or even in the opposite direction over time compared to the effect in the beginning. On the flip side, if opinions can persistently predict future returns, then we would find a drift coefficient over time.

The two images on the left of Figure 2 display the estimated coefficients on non-trading-hour opinions for CARs from a week to half a year. For pessimistic opinions published during non-trading hours, the magnitudes of these coefficients are similar, but they are imprecisely estimated for longer holding periods. For positive opinions published during non-trading hours, CARs exhibit a slightly upward drift. The two images on the right also reveal that the estimated coefficients for trading-hour opinions tend to increase with the length of the holding period. One possible explanation for these drifts is that trading-hour opinions and non-trading-hour positive opinions have relatively persistent effects as indicated in columns 5–9 of Table 3. Therefore, we can expect that effects predicted by these opinions will accrue when examining CARs over longer holding periods.

4. Explanations of results

4.1 Sentiment or fundamental

To check whether *Guba* opinions contain value-relevant information or just capture sentiment, we conduct two exercises. First, Baker and Wurgler (2006) show that sentiment has a stronger impact on stocks that are difficult to evaluate, such as small stocks. It suggests that analyzing returns across different size groups would be an effective way to test whether *Guba* opinions proxy for investor sentiment.

Table 5 displays the estimated results for stocks that are divided into 5 groups based on their sizes, as determined by the market capitalization of firms as of last June. All columns include the same controls as those in column 5 of Table 3. Panel A shows that non-trading-hour opinions have a stronger predictive power for returns for big stocks than small stocks. We observe a clear increasing pattern in the magnitude of coefficients of $NPess_{it}$ and $NPos_{it}$ across these columns. The last column reports the *p*-values of the F-test checking whether the coefficients are jointly equal to 0 among the previous columns. All these *p* values are 0. For trading-hour opinions, the coefficients of $TPess_{it}$ and $TPos_{it}$ are both significant, but their magnitude seems unrelated to firm size.

In Panels B and C, we decompose daily returns into overnight and daytime returns. Panel B shows that the effects of non-trading-hour opinions on overnight returns increase with firm size. In Panel C, we use daytime returns as dependent variables in order to examine whether daytime return reversals vary by firm size. We find that the magnitude of non-trading-hour opinions increases with group size. The last row reports the percent of returns that reverse during the next trading day relative to the overnight return fluctuation. We find that this proportion does not vary with firm size.[22] Overall, these findings suggest that individual investors, who are the main users of the *Guba* forum, tend to trade big stocks more during market opening hours. By contrast, institutional investors, who may possess valuable information about big stocks, trade in the opposite direction on the next trading day.

These results show that opinions have stronger effects on big firms, suggesting that the predictability of *Guba* opinions on returns is not solely driven by investor sentiment. As we will show in the following text, *Guba* articles contain no value-relevant information for small firms, indicating that the price fluctuations of small stocks are likely driven by investor sentiment.

We then test whether *Guba* opinions contain valuable information that is not reflected in the stock price at the time of their publication. If the amount of valuable information contained in *Guba* articles differs between small and big firms, then we may also find that the return predictability of opinions varies with firm size. We follow Chen et al. (2014) and

[22] We also categorize stocks into five groups based on book-to-market ratio, number of analyst coverage, and R&D expenditures. Table A4 in the Appendix shows that non-trading-hour opinions have a relatively low ability to predict return for high-growth firms (low B/M firms) and distressed firms (high B/M firms). These two types of stocks are difficult to value (Baker and Wurgler, 2006). Panels B and C do not reveal a clear relationship between the return predictability of *Guba* opinions and analyst coverage or R&D expenditure.

check whether *Guba* opinions can predict earnings surprise. We obtain yearly data on the profits and earnings forecasted by analysts for 2,987 stocks between 2008 to 2021. We measure earnings surprise using two variables: (1) the difference between reported earnings per share (EPS) and the average EPS forecasted by analysts, and (2) the difference between the reported profits and the average profits forecasted by analysts.[23] We also construct two dummies, each of which takes a value of 1 if each measure of the aforementioned earnings surprise is greater than 0. We regress earnings surprise on trading-hour and non-trading-hour *Guba* opinions from the previous day. We control for the number of reading times and comments, a dummy indicating for newspaper coverage, the number of pessimistic and positive articles, and the total number of newspaper articles related to the firm. All these control variables are measured on the same day as our opinion variables. We also control for firm fixed effects. Standard errors are clustered at the firm level.

We examine whether the effect of *Guba* opinions on earnings surprise varies with firm size. To save statistical power, we categorize stocks into two size groups. Table 6 reveals that *Guba* opinions have a strong predictive power for big firms but not for small firms. For example, Panel A shows that the estimated coefficients on opinions for small firms are mostly small and insignificant, which is consistent with the fact that small firms have higher information uncertainty and are difficult to analyze (Chan et al., 1985; Loh and Stulz, 2018). By comparison, column 1 of Panel B shows that, for big firms, their earnings surprise will decrease (increase) by 0.03 RMB (0.02 RMB) if the proportion of non-trading-hour pessimistic (positive) opinions increases from the 25th to the 75th percentiles. Column 2 of Panel B shows that moving non-trading-hour pessimistic (positive) opinions from the 25th to the 75th percentiles will decrease (increase) the probability of a positive earnings surprise by 1% (2%) for big firms. Additionally, even for big firms, the effects of trading-hour opinions are mostly insignificant. These results suggest that value-relevant information is mainly contained in articles that are published during non-trading hours. Moreover, this valuable information is about big stocks.[24] Combined with our previous finding in Table 5, these results help us understand why non-trading-hour opinions have a

23 We measure such difference in 1 million RMB.

24 Our finding that online stock forum contains some relevant information is consistent with Chen et al. (2014), who find that investors can exploit value-relevant information through discussions and interactions in the US online forums, a phenomenon known as the "wisdom of crowds."

greater ability to predict returns for big firms. Moreover, given that *Guba* opinions contain little value-relevant information for small firms, the influence of opinions on these stocks is primarily driven by the sentiment of noise investors.

4.2 Trading hour versus non-trading hour

As we illustrate above, non-trading-hour opinions seem to contain more value-relevant information. One possible reason is that government and firms disclose a substantial amount of material public information after the market closes (Barclay and Hendershott, 2003; Santosh, 2016). For example, the People's Bank of China (PBC, the central bank) buys and sells bonds in the open market from 9:00 AM to 9:20 AM when needed and releases this information at 9:20 AM. When the PBC plans to adjust its interest rates, such adjustment is often implemented over the weekend. Chinese listed companies follow a similar rule. Specifically, the listed companies often disclose important information, known as "major events" about their operations after the stock market closes to avoid stock price fluctuation. A major event is defined as an event that may greatly affect share prices, such as the appointment of a new CEO, change in dividend policies, acquisition, etc. Table A2 in the Appendix lists these events and the frequency at which each event occurs during our sample period. The three most common events are the firms' future earnings and profit forecasts, share repurchases, and CEO changes. Figure A3 in the Appendix displays the daily time-series proportion of firms announcing major events. The figure clearly illustrates variations in this proportion throughout the sample period. Figure 3 illustrates the distribution of hours of press conferences where companies announce most of their major events. Our data cover the years 2012 to 2021.[25] The above figure shows the start time of these conferences, while the bottom figure shows the end time. The beginning time of these press conferences stacks at 3:00 PM, indicating that the relevant information is mostly published after the market closes. Approximately 85% of the conferences are held after the market close.

Given that firms may announce major events, which contain material information after the market closes, it is worth exploring whether the predictability of *Guba* opinions is caused by these events. In our exercise, we control for these major events and explore

25 We cannot obtain such data from 2012.

whether our main findings hold. Based on Equation (1), we further control a dummy indicating a firm's announcements of major events on day *t*. We control the same covariates shown in column 5 of Table 3. Panel A of Table 7 reports the results for ARs from days *t*+1 to *t*+5. Our results barely change. As mentioned in Section 2.1, companies typically utilize their official accounts to post material information simultaneously on *Guba* forum on major-event days, implying that this material information is available through the forum. However, after controlling for this information, we still observe the strong predictive ability of *Guba* opinions on stock returns in the subsequent days. This result suggests that the predictive power of *Guba* opinions is not driven by these major events or the information they contain. Our results are robust when we use the two-step regression method proposed by García (2013). Specifically, in the first step, we regress each of our opinion variables on a dummy indicating for the major-event day, stock returns from days *t*−5 to *t*, opinion variables from days *t*−5 to *t*, various controls, and firm and month-by-year fixed effects. In the second step, we replace the opinions with the residuals obtained from the first step as our explanatory variables. Our results are robust to this exercise (Table A5 in the Appendix).

Whether the opinions published on major-event days have a greater impact on returns? To answer this question, we compare the effects of *Guba* opinions published on major-event and regular days by using the same controls and fixed effects shown in column 5 of Table 3. Panel B of Table 7 reports the coefficients that measure the effect of *Guba* opinions that are published on major event days on stock returns. We find clear evidence supporting the return predictability of trading- and non-trading-hour opinions.

Panel C of Table 7 presents the coefficients that measure the effect of *Guba* opinions that are published during regular days on stock returns. Those opinions can also move stock returns on the following days, thereby corroborating our previous finding that the return predictability of *Guba* opinions does not solely work through the firms' announcements of major events. However, the effects of opinions published on regular days are smaller: the magnitude of coefficients in Panel C is notably smaller than that in Panel B, particularly for non-trading-hour opinions. The effects of opinions published on major-event days are approximately 1.5–2 times greater than those of opinions published on regular days. The bottom rows compare the coefficients in Panels B and C and indicate that, except for trading-hour positive opinions, most of the differences are statistically significant. In sum, the predictability of *Guba* opinions on stock returns is significantly larger for opinions published on days when firms announce major events.

4.3 Investor attention

We have observed that the predictability of opinions on returns is not driven by firms' announcements of major events and that non-trading-hour opinions have a greater impact on returns when they are published on major-event days. To understand these results, we examine investor attention and demonstrate that the opinions published on major-event days tend to attract more attention. A possible reason for these results, as we will describe below, is that the announcement of such events exposes investors to higher levels of uncertainty.

We use reading times and number of comments to measure investor attention. Given that these attention variables are counted for each article, we can categorize them into trading- and non-trading-hour categories. These opinions may be read or commented on several days after their publication. We do not exclude "late" readings and comments because they can help us understand the sustained investor attention to these opinions.[26] Given that our aim is to compare investor attention associated with trading- and non-trading-hour articles, we do not distinguish articles by their tones. Specifically, our two explanatory variables are the number of articles published during trading- and non-trading hours. Noteworthily, the coefficients on the number of articles cannot be interpreted as "causal" because having more reads and comments may attract investors to express their opinions by publishing more articles.

We regress reading times/number of comments on the number of trading- and non-trading-hour articles using the same controls and fixed effects as those shown in column 5 of Table 3. We find that non-trading-hour articles receive more investor attention on average. Specifically, column 1 of Table 8 shows that the point estimate of "# of non-trading-hour articles" is 1,089.6 (t=7.27), implying that a non-trading-hour article receives

26 We used a computer program to collect data from the *Guba* forum in 2015. This program ran on 20 computers for almost two months, resulting in a sample of 50 stocks and nearly 65,000 observations. One advantage of our collected data is that they include the time when posts are published and when comments are made. We consider those comments made days after a post was published as "late comments," which account for approximately 17% of all comments in our sample. We also compare the reading times of posts published on the day of the data collection with those of posts that were published several days ago and find that the latter were read 9% more compared with the former.

1,089.6 reads on average. Column 2 shows that the coefficient on "# of trading-hour articles" is 492.3 (t=11.13), indicating that a trading-hour article receives 492.3 reads on average. In column 3, we aggregate the number of reading times for articles published during trading- and non-trading hours and perform a horse race by simultaneously regressing this aggregated variable on the number of trading- and non-trading-hour articles. The coefficient of the number of non-trading-hour articles is almost twice as large as that of the number of trading-hour articles. In columns 5–7, we use the number of comments as the dependent variable and find a more significant and stronger correlation between comments and non-trading-hour articles than between comments and trading-hour articles. Specifically, a non-trading-hour article receives 2.9 comments on average, while a trading-hour article only receives 1.4 comments on average.

In columns 4 and 8, we further add a dummy indicating firms' announcements of major events and its interaction with the number of trading- or non-trading-hour articles. We find that the estimated coefficients on these interactions are positive and significant for the number of non-trading-hour articles but negative for the number of trading-hour articles. This result suggests that investors pay more attention to non-trading-hour articles when firms announce major events. This evidence provides a possible explanation for why non-trading-hour opinions have a greater impact on returns for those published on major-event days as shown in Table 7. Our finding is in line with Clarke et al. (2020), who show that fake news attracts more investor attention and affects stock prices in a US stock discussion online forum.

We conjecture that the release of major events exposes inexperienced retail investors to high uncertainty. Because these investors lack the financial knowledge to interpret the information, they seek advice from others through online forums to help them reduce the risk. To test this argument, we conduct an event study by comparing the volatility of stocks around the days when firms announce major events (i.e., day = 0). For each day, we calculate the standard deviation of a stock price over the 5-, 10-, and 20-day windows before day t and consider it as the standard deviation of this stock on day t.[27] We then roll windows to calculate the volatility on each day. Figure 4 presents the volatility 20 days

27 Similar to the uncertainty at the market level, Frankel et al. (2006) and Loh and Stulz (2018) find that firm-level uncertainty is an important factor affecting investor behavior.

before and after the announcement of a major event. The volatility increases dramatically on the next trading day following the announcement of major events, thereby implying that the announcement of these events exposes investors to a high level of uncertainty.

We then examine whether investors pay more attention to opinions when they are exposed to high uncertainty. We run regressions similar to that shown in column 4 of Table 8 but replace the dummy and its interactions with volatility and its interactions with the number of trading- and non-trading-hour articles. We use the 5-day rolling window standard deviation of the stocks to measure uncertainty.[28] Column 2 of Table 9 shows that, when stock volatility increases from the 25th to the 75th percentiles, the non-trading-hour articles receive 439 ($436.5 = 18{,}977.1 \times 0.023$) more times of reads.[29] This effect is statistically different from 0 at the 1% level. By comparison, trading-hour articles are read fewer times when stock volatility is high. Similarly, column 4 shows that, when stock volatility increases from the 25th to the 75th percentile, the non-trading-hour articles significantly receive 0.86 ($0.87 = 37.6 \times 0.023$) more comments. Again, trading-hour articles receive fewer comments when stock volatility is high. Our findings reveal that investors are more attracted to articles on the online forum when they are exposed to high uncertainty. Our results help us to understand why non-trading-hour opinions published on major-event days have a higher ability to predict stock returns: they receive more investor attention. Our findings are consistent with Frankel et al. (2006) and Loh and Stulz (2011, 2018), who show that investors rely on the opinions of other people when faced with high uncertainty.

Overall, firms' announcements of major events after the market closes increase the volatility of stock prices. Retail investors may struggle to interpret this information and rely on *Guba* to seek advice from other people. Investors who exhibit herd behaviors will create high price pressure at the opening of the market (De Bondt and Thaler, 1985; Scharfstein and Stein, 1990). This logic also helps explain why daytime return reversals are only observed following the publication of non-trading-hour opinions.

---

28 Using the volatility calculated in the rolling window of the past 5, 10, and 20 days can obtain similar results.

29 The 25th and 75th percentiles of volatility are 0.009 and 0.032, respectively.

5. Further analysis and robustness

5.1 Are return predictability of *Guba* opinions related to market conditions?

Determining the return predictability of *Guba* opinions across different market conditions is valuable. First, the market environment creates a certain structure in which the actions of investors respond differently to online opinions during good and bad times (Guidolin and Timmermann, 2005; Kim and Nofsinger, 2007; Necker and Ziegelmeyer, 2016). For example, Kim and Nofsinger (2007) show that investors exhibit some striking differences in investing behavior between the bull and the bear market. Cujean and Hasler (2017) indicate that investors interpret the same news differently depending on the present economic conditions. Second, the biased expectations of investors may be amplified in bad times. For example, Veronesi (1999) shows that investors tend to overreact to bad news during good periods and underreact to good news during bad periods. An examination of the effects of *Guba* opinions on returns in good and bad periods is necessary.

We use the approach of Bry and Boschan (1971) to identify the peaks and troughs of the Chinese stock market in our sample period.[30] Specifically, the monthly stock market index (CSI 300 Index) is used to identify the local maximum points (peaks) where the market index is higher than that in the neighboring 5-month window on both sides.[31] In a similar vein, local minimum points (troughs) are those where the market index is lower than its neighboring 5-month window. Specifically, we identify 4 peaks (2009 July 31, 2015 May 29, 2018 January 31, and 2021 January 29) and 4 troughs (2008 October 31, 2012 November 30, 2016 February 29, and 2018 December 28) in our sample period.[32] Accordingly, we can obtain the start and end dates of good and bad times and divide the sample into 5 bad times and 4 good times. Figure A4 shows the peaks, troughs, and bad and good times.

We examine abnormal returns from t+1 to t+5. We use Equation (1) to regress each of the above variables on four *Guba* opinion variables, controls, and year-by-month fixed effects separately in good and bad periods, respectively. Figure 5 illustrates the estimated

30 This approach has been widely adopted in the literature (e.g., Pagan and Sossounov, 2003; Gonzalez et al., 2005; Yan et al., 2007).
31 The CSI 300 Index is one of the most widely influential stock market indices in China. This value-weighted index comprises 300 stocks selected from the Shanghai and Shenzhen Stock Exchanges.
32 Our data show that the minimum length of a complete cycle (i.e., from peak to peak or from trough to trough) is 12 months, and the minimum length of a phase (i.e., from peak to trough or from trough to peak) is 5 months.

coefficients and 95% confidential intervals of the four *Guba* opinion variables. The left figure shows the estimated results for bad times and the right figure shows those for good times. *Guba* opinions can significantly affect stock returns in both good and bad times. Moreover, positive opinions tend to have higher effects on stock returns during good times, implying that investors tend to overestimate good opinions when the stock market surges. Our results are different from Barberis et al. (1998) who show that following a string of positive shocks in good times, the responses of investors are relatively weak since such positive shocks have been anticipated. The difference between our findings and Barberis et al. (1998) may be that we examine a context that comprises a high percentage of retail investors who may lack financial knowledge and respond to these shocks with overzealous investment actions.

5.2 Can investors earn profits by following *Guba* opinions?

An interesting question is whether investors have the opportunity to arbitrage by using a strategy that is leveraged on the *Guba* opinions. We answer this question by incorporating transaction costs into portfolio returns. For ease of narration, we refer to the strategy that involves the analysis of *Guba* opinions to buy stocks as "opinion strategy."

The cost of an A-share transaction in China mainly has three components, namely, commission, stamp tax, and slippage (Leippold et al., 2021). The commission fee is 0.03% and is collected for both buying and selling shares. The stamp tax is 0.1% and is collected unilaterally from sellers. Slippage is 0.06% and is collected for both buying and selling shares.[33] The overall cost for a round-trip transaction is 0.28%.

First, we notice that the non-trading-hour pessimistic opinions have the highest effects on abnormal returns on the next day. Column 5 of Table 3 shows that the estimated coefficient is −0.65, which implies that an investor can earn an excess return of 0.065% by short-selling a stock that receives 100 percent of negative comments. It is evident that it is impossible to earn profits via the opinion strategy if transaction costs are considered.

Second, we consider portfolio construction to analyze whether an investor can earn profit by using the opinion strategy in the long run. We only consider non-trading-hour *Guba* opinions given that they have a higher impact on future returns. The portfolio is constructed

33 The slippage fee is only charged for the transaction of stocks listed in the Shanghai Stock Exchange.

as follows. The length of a period over which the portfolio is held is denoted by *K*. To avoid microstructure effects, one should wait for a certain period before re-implementing his/her trading strategy again. We denote this waiting period by *S*. Therefore, vector (*K*, *S*) describes an opinion strategy. We consider *K*=60, 120. We let *S*=5, 20, with the implication that an investor implements the portfolio strategy every week or month. We use the equal-weighted strategy to construct the portfolios. Results from the value-weighted strategy are quantitively similar.

We follow the literature to consider the strategy by calculating the profits of past "losers" versus past "winners" (Grinblatt and Moskowitz, 2004; Jegadeesh and Titman, 2001). Many studies define "winners" as those firms having the top 10% ranking-period performance, and "losers" as those having the bottom 10% ranking-period performance. Following this convention, for non-trading-hour pessimistic opinions, we define stocks as "winners" if the proportion of non-trading-hour pessimistic opinions of stocks belongs to the top 10% and "losers" if the proportion belongs to the bottom 10%. For non-trading-hour positive opinions, we define the "winners" and "losers" in a reverse manner. Column 1 of Table A6 shows that only the pessimistic opinion strategies (120, 5) can yield a profit of 0.1%. However, when we consider the transaction cost (0.28%), this strategy cannot earn money either. Moreover, short selling is strictly controlled in China.[34] Korajczyk and Sadka (2004) suggest that allowing short selling may violate the up-tick rule. In columns (2) and (4) of Table A6, we restrict short-selling and limit our analysis to "winners." All returns are negative, implying that investors cannot earn profits in a market.

5.3 Robustness

We conduct four exercises to show that our results are robust. First, one may be concerned about the autocorrelation of opinions and the correlation between stock return on date *t* and *Guba* opinions on date *t*. To address this issue, we follow García (2013) to regress each of our opinion variables on day *t* on stock returns from days *t*–5 to *t* and controls.[35] We

34 Although China has gradually introduced short selling by letting insurers lend securities since 2010, but short selling has been strictly restrained since the stock market crashed in 2015. The size of the securities lending remains small. For example, the outstanding value of securities lending was less than 9% of the outstanding value for margin loans in 2020.

35 Specifically, our specification is: $opinion_t = \gamma_0 + \gamma_1 Ret_t + \gamma_2 L(Ret_t) + \gamma_3 L(opinion_t) + \mathbf{X}_{it}\boldsymbol{\delta} + \mu_i + \mu_{my} + \epsilon_{it}$, where $opinion_t$ represents one of our opinion measures ($Pess_{it}$, $NPos_{it}$, $TPess_{it}$, $TPos_{it}$). Ret_t denotes the stock abnormal returns on day *t*. $L(Ret_t)$ represents lagged terms of stock abnormal returns from days *t*–5 to *t*–1. $L(opinion_t)$ represents lagged terms of the corresponding opinion variables from days *t*–5 to *t*–1. $\mathbf{X}_{it}$ denotes a set of controls that are similar to column 5 of Table 3 but without previous ARs and CARs.

replace our key explanatory variables, namely, the four opinion variables in Equation (1), with the residuals from the above model. Panel A of Table 10 shows that the coefficients on our opinion variables are similar to those in Table 3.

Second, while we have accounted for firm fixed effects in our analysis, the potential impact of firms' time-varying characteristics also warrants consideration. To address this concern, we include additional controls of lagged firm size (measured in logarithm), book-to-market ratio, and trading volume, which vary by day and can be included in our model. The robustness of our results is demonstrated in Panel B of Table 10.

Third, given that certain industries may be more likely to be discussed and focused by investors, a potential systematic link may exist between these industries and our measure of *Guba* opinion variables. This may lead to biased estimates of *Guba* opinions if the latter is systematically correlated with the unobserved heterogeneity of the industry. To address this issue, we include industry-by-year fixed effects to allow industry heterogeneity to vary by year. Panel C of Table 10 shows that our main findings are unchanged.

Finally, we conduct a placebo test where we randomly assign abnormal returns from days t+1 to t+5 on *Guba* opinions. We re-run the regression of our main analysis and save the estimated coefficients of the four *Guba* opinion variables. We repeat the above exercise 500 times. Table A7 in the Appendix reports the average coefficient over 500 repetitions and the percentage of coefficients that are significant at the 5% level. Consistent with our expectations, less than 5% of the coefficients are significant at the 5% level, which indicates that our findings regarding the return predictability of *Guba* opinions are not generated by random noise.

6. Conclusion

In this study, we employ data from one of the most popular online stock discussion forums in China to investigate the effects of stock opinions published during trading periods and after the market closes on stock returns. We find that both trading- and non-trading-hour opinions can predict future stock returns. Non-trading-hour opinions have a more pronounced effect than trading-hour opinions. However, a substantial portion of the returns affected by non-trading-hour opinions reverses during the next trading period. These results suggest that individual investors, who are the main users of the online forum, tend to trade near the opening of the market, while institutional investors are likely to trade in the opposite direction during the trading periods.

We find that non-trading-hour opinions have a higher impact on stock returns for large stocks. This is because articles published after the market closes offer value-relevant information about large firms. Given that no value-relevant information was released about small firms, the effects of opinions for small firms are primarily driven by the sentiment of noise investors. In addition, we show that non-trading-hour opinions published on days when firms announce major events have a higher ability to predict returns. These opinions also attract more investor attention. A possible explanation is that firms' disclosure of major events exposes retail investors to a higher level of risk, leading them to seek advice from other investors on online forums.

An increasing number of studies have investigated the impact of social media, online forums, and online social networks on stock returns. In our research, we contribute to this field by presenting new evidence regarding the different effects of trading- and non-trading-hour opinions. Given that the online forum is more favored by individual investors, our findings shed light on the sources of return reversals during trading hours following the price pressure that occurs at the market's opening. While we have addressed some significant questions, our findings also raise additional inquiries that warrant further investigation. For instance, future studies could explore whether there are organized behaviors that influence the timing of opinion publication and how these behaviors may influence stock returns.

References

Akbas, F., Boehmer, E., Jiang, C., and Koch, P. D. (2022). Overnight returns, daytime reversals, and future stock returns. *Journal of Financial Economics* 145: 850–875.

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

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

Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market, *Journal of Economic Perspectives* 21: 129–151.

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

Barclay, M. and Hendershott, T. (2003). Price discovery and trading after hours. *Review of Financial Studies* 16 (4): 1041–1073.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics* 49: 307–343.
Berkman, H., Koch, P., Tuttle, L., and Zhang, Y. (2012). Paying attention: overnight returns and the hidden cost of buying at the open. *Journal of Financial and Quantitative Analysis* 47(4): 715–741.
Bry, G. and Boschan, C. (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. National Bureau of Economic Research, Inc: New York.
Chan, K.C., Chen, N.F., and Hsieh, D.A. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics*14(3): 451–471.
Chen, H., De, P., Hu, Y.J., and Hwang, B.H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27(5): 1367–1403.
Clarke, J., Chen, H., Du, D., and Hu Y.J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research* 32(1):35–52.
Cujean, J. and Hasler, M. (2017). Why does return predictability concentrate in bad times? *Journal of Finance* 72(6): 2717–2758.
Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *Journal of Finance* 66(5): 1461–1499.
Das, S. and Chen, M.Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science* 53(9): 1375–1388.
De Bondt, W.F., and Thaler, R. (1985). Does the stock market overreact? *Journal of Finance*, 40(3): 793–805.
Epstein, L.G. and Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory* 113(1): 1–31.
Epstein, L.G. and Schneider, M. (2008). Ambiguity, information quality, and asset pricing. *Journal of Finance* 63 (1):197–228.
Fama E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22.
Frankel, R., Kothari, S.P., and Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41: 29–54.
García, D. (2013). Sentiment during recessions. *Journal of Finance* 68(3): 1267–1300.
Gonzalez, L., Powell, J.G., Shi, J., and Wilson, A. (2005). Two centuries of bull and bear market cycles. *International Review of Economics & Finance* 14(4): 469–486.
Grinblatt, M. and Moskowitz, T.J. (2004). Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics*, 71(3): 541–579.
Guidolin, M. and Timmermann, A. (2005). Economic implications of bull and bear regimes in UK stock and bond returns. *Economic Journal* 115(500): 111–143.
Hirshleifer, D. (2020). Presidential address: Social transmission bias in economics and finance. *Journal of Finance* 75(4): 1779–1831.
Huang, J. (2018). The customer knows best: The investment value of consumer opinions. *Journal of Financial Economics* 128(1): 164–182.
Huang, Y., Qiu, H., and Wu, Z. (2016). Local bias in investor attention: Evidence from China's internet stock message boards. *Journal of Empirical Finance* 38: 338–354.
Jegadeesh, N. and Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance*, 56(2): 699–720.

Jiang, L., Liu, J., and Yang, B. (2018). Communication and comovement: Evidence from online stock forums. *Financial Management* 48(3): 805–847.
Kelley, E.K. and Tetlock, P.C. (2013). How wise are crowds? Insights from retail orders and stock returns. *Journal of Finance* 68(3): 1229–1265.
Kim, K.A. and Nofsinger, J.R. (2007). The behavior of Japanese individual investors during bull and bear markets. *Journal of Behavioral Finance* 8(3): 138–153.
Kim, S.-H. and Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization* 107: 708–729.
Korajczyk, R. and Sadka, R. (2004). Are momentum profits robust to trading costs? *Journal of Finance* 59(3): 1039–1082.
Leippold, M., Wang, Q., and Zhou, W. (2021). Machine learning in the Chinese stock market. *Journal of Financial Economics*. *https://doi.org/10.1016/j.jfineco.2021.08.017*.
Leung, H. and Ton, T. (2015). The impact of internet stock message boards on cross-sectional returns of small-capitalization stocks. *Journal of Banking & Finance* 55: 37–55.
Loh, R.K. and Stulz, R.M. (2011). When are analyst recommendation changes influential? *Review of Financial Studies* 24: 593–627.
Loh, R.K. and Stulz, R.M. (2018). Is sell-side research more valuable in bad times? *Journal of Finance* 73(3): 959–1013.
Lou, D., Polk, C., and Skouras, S. (2019). A tug of war: Overnight versus intraday expected returns. *Journal of Financial Economics* 134(1): 192–213.
Lou, D., Polk, C., and Skouras, S. (2022). The day destroys the night, night extends the day: A clientele perspective on equity premium variation. LSE working paper.
Necker, S. and Ziegelmeyer, M. (2016). Household risk taking after the financial crisis. *Quarterly Review of Economics and Finance* 59: 141–160.
Pagan, A.R. and Sossounov, K.A. (2003). A simple framework for analysing bull and bear markets. *Journal of Applied Econometrics* 18(1): 23–46.
Renault, T. (2017). Intraday online investor sentiment and return patterns in the US stock market. *Journal of Banking & Finance* 84: 25–40.
Santosh, S. (2016). The Path of price discovery: Trade time vs. clock time. Working Paper. University of Colorado at Boulder.
Scharfstein, D. S. and Stein, J. C. (1990). Herd behavior and investment. *American Economic Review*: 465–479.
SEC. (2012). Investment adviser use of social media. *National Examination Risk Alert* 2:1–7.
Tumarkin, R. and Whitelaw, R.F. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*, 57(3):41–51.
Veronesi, P. (1999). Stock market overreactions to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* 12(5): 975–1007.
Williams, C. (2009). Asymmetric responses to good and bad news: An empirical case for ambiguity. Working paper, University of Michigan.
Yan, W., Powell, J.G., Shi, J., and Xu, W. (2007). Chinese stock market cyclical regimes: 1991–2006. *Economics Letters* 97(3): 235–239.
You, J., Zhang, B., and Zhang, L. (2017). Who captures the power of the pen? *Review of Financial Studies* 31(1): 43–9.

Table 1. Summary Statistics of Variables of *Guba* Data

	Non-trading hours				Trading hours				(1)−(5)
	Mean	Standard deviation	Min	Max	Mean	Standard deviation	Min	Max	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Proportion of pessimistic articles	0.174	0.185	0	1	0.272	0.169	0	1	−0.098***
Proportion of positive articles	0.345	0.241	0	1	0.270	0.173	0	1	0.075***
Number of pessimistic articles	2.771	7.208	0	4,525	8.108	16.342	0	2,800	−5.337***
Number of positive articles	5.212	11.915	0	3,700	7.494	14.040	0	3,200	−2.282***
Number of articles	16.647	38.235	1	18,000	29.818	57.925	1	78,00	−13.172***
Number of reading times	22,000	230,000	1	400,000,000	25,000	59,000	1	23,000,000	−3,000***
Number of comments	41.562	178.615	0	89,000	50.103	357.809	0	600,000	−8.541***
Average reading times per article	1,400	27,000	1	50,000,000	1,000	1,800	1	1,100,000	400***
Average comments per article	2.430	15.690	0	22,000	1.951	11.127	0	17,000	0.479***
Receive any reads (yes=1)	1	0	1	1	1	0	1	1	0
Receive any comments (yes=1)	0.868	0.339	0	1	0.953	0.213	0	1	−0.085***
Observations		3,422,599				3,422,599			

Notes: This table presents the summary statistics of the variables during trading and non-trading hours using data from *Guba*. The sample period is from 2008 to 2021. Proportion of pessimistic articles and positive articles are proportions of pessimistic and positive articles on stock *i* posted on day *t*. Number of pessimistic and positive articles are the number of pessimistic and positive articles on stock *i* posted on day *t*. Number of reading times is the number of reading times received by stock-*i* related articles. Number of comments is the number of comments received by stock-*i* related articles. Receive any comments is a dummy that takes a value of 1 if at least one article on stock *i* receives comments on day *t*, and 0 otherwise. Receive any reading is a dummy that takes a value of 1 if at least one article on stock *i* is read on day *t*, and 0 otherwise.

Table 2. Summary Statistics of Stock Returns and Other Variables

	Observations	Mean	Standard deviation	Min	Max
	(1)	(2)	(3)	(4)	(5)
Panel A. Abnormal returns and CARs					
$Ret_{i,t+1}$	3,422,599	−0.0006	0.0234	−0.1597	0.2528
$Ret_{i,t+2}$	3,164,132	−0.0006	0.0234	−0.1597	0.2528
$Ret_{i,t+3}$	3,193,154	−0.0006	0.0233	−0.1597	0.2528
$Ret_{i,t+4}$	3,182,779	−0.0006	0.0233	−0.1597	0.2528
$Ret_{i,t+5}$	3,174,746	−0.0006	0.0232	−0.1588	0.2528
$CAR_{i,t+5}$	3,391,087	−0.0014	0.0521	−0.5511	0.6574
$CAR_{i,t+10}$	3,370,552	−0.0027	0.0715	−0.9704	1.0574
$CAR_{i,t+20}$	3,339,209	−0.0049	0.0986	−1.2251	1.4513
$CAR_{i,t+60}$	3,239,415	−0.0119	0.1624	−1.2963	2.0144
$CAR_{i,t+120}$	3,116,237	−0.0217	0.2243	−1.5537	2.1142
Ret_{it}	3,422,599	−0.0364	0.3124	−2.0145	2.3301
$Ret_{i,t-1}$	3,422,599	−0.0005	0.0235	−0.1628	0.2528
$Ret_{i,t-2}$	3,422,599	−0.0004	0.0237	−0.1731	0.2528
$CAR_{i,t-60,t-3}$	3,422,599	−0.0003	0.0238	−0.1731	0.2528
Panel B. Firms' characteristics and newspaper variables					
Log (market value)	3,422,599	22.6158	1.0147	19.3487	29.2543
Book-to-market ratio	3,422,599	0.4567	0.2997	0	5.5556
Revenue/market value	3,422,599	0.4393	0.8537	0	56.4531
Profit/market value	3,422,599	0.0251	0.0418	−1.4070	0.8671
R&D/market value	3,422,599	0.0152	0.0288	0	2.8049
Trading volume	3,422,599	2.1310	4.6530	0.0016	343.7512
INews	2,974,797	0.0463	0.2101	0	1
NegNews	2,974,797	0.0260	0.2070	0	36
PosNews	2,974,797	0.0263	0.1941	0	15
Newsnum	2,974,797	0.0655	0.3749	0	36
Analyst coverage	3,422,599	0.2312	0.6953	0	6
Major event	3,422,599	0.0887	0.2843	0	1
Earnings surprise: EPS	21,548	−0.0011	0.0355	−11.2674	6.82
Earnings surprise: positive EPS	21,548	0.0007	0.0264	0	1
Earnings surprise: Netprofit	21,291	−0.0563	7.4569	−5,500	253.4983
Earnings surprise: positive Netprofit	21,291	0.0008	0.0274	0	1

Notes: This table presents the summary statistics of returns and control variables. The sample period is from 2008 to 2021. Trading volume is measured in 100 million RMB. *INews* takes the value of 1 if firm i is covered by any newspapers on day *t* and 0 otherwise. *NegNews*, *PosNews*, and *Newsnum* are the number of firm-*i* related pessimistic, positive articles, and the total number of newspaper articles published on day *t*. Analyst coverage is the number of analysts who cover the firm. *Report* and *Major event* are two dummies that indicate whether a firm publishes a financial report on day *t* and whether a firm discloses major events on day *t*. "Earnings surprise: EPS" is the difference between the reported EPS and the average EPS forecasted by analysts. "Earnings surprise: Netprofit" is the difference between the reported profits and the average profits forecasted by analysts in 100 million RMB. "Earnings surprise: positive EPS" and "Earnings surprise: PosNetprofit" are two dummies that indicate whether the announced earnings per share and the announced net profit of firm *i* on day *t* are higher than the prior forecasts of analysts.

Table 3. Effects of *Guba* Opinions on Abnormal Returns in the Subsequent Five Days

	$Ret_{i,t+1}$ (1)	$Ret_{i,t+1}$ (2)	$Ret_{i,t+1}$ (3)	$Ret_{i,t+1}$ (4)	$Ret_{i,t+1}$ (5)	$Ret_{i,t+2}$ (6)	$Ret_{i,t+3}$ (7)	$Ret_{i,t+4}$ (8)	$Ret_{i,t+5}$ (9)
(a) $NPess_{it}$	−0.752***		−0.730***	−0.651***	−0.651***	0.113	0.073	0.015	−0.084
	(−10.61)		(−10.32)	(−8.95)	(−8.95)	(1.48)	(0.92)	(0.19)	(−1.07)
(b) $NPos_{it}$	1.060***		1.041***	0.980***	0.979***	0.333***	0.193***	0.187***	0.180***
	(18.27)		(17.95)	(16.46)	(16.45)	(5.29)	(3.04)	(3.01)	(2.89)
(c) $TPess_{it}$		−0.496***	−0.452***	−0.492***	−0.493***	−0.519***	−0.298***	−0.332***	−0.240***
		(−6.53)	(−5.96)	(−6.13)	(−6.14)	(−5.93)	(−3.38)	(−3.87)	(−2.76)
(d) $TPos_{it}$		0.487***	0.404***	0.414***	0.413***	0.390***	0.555***	0.415***	0.517***
		(6.50)	(5.41)	(5.37)	(5.36)	(4.57)	(6.59)	(4.85)	(5.90)
$INews_{it}$				0.120***	0.132***	−0.045	−0.117***	−0.075*	−0.045
				(2.84)	(3.13)	(−1.05)	(−2.82)	(−1.80)	(−1.07)
$NegNews_{it}$					−0.002	1.000	0.012*	−0.020***	−0.004
					(−0.79)	(0.09)	(1.94)	(−5.72)	(−0.71)
$PosNews_{it}$					0.013***	0.007	−0.015*	0.039***	−1.000
					(2.77)	(0.77)	(−1.69)	(7.45)	(−0.08)
$Newsnum_{it}$					−0.055***	−0.02	−0.030*	0.007	0.060***
					(−3.79)	(−1.14)	(−1.88)	(0.55)	(3.44)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,422,599	3,422,599	3,422,599	2,974,797	2,974,797	2,743,955	2,771,363	2,763,199	2,756,885
R^2	0.004	0.004	0.004	0.004	0.004	0.002	0.003	0.003	0.003
p-value of (a)+(b)=0	0.003		0.003	0.002	0.002	0.000	0.024	0.084	0.410
p-value of (c)+(d)=0		0.938	0.703	0.188	0.546	0.383	0.074	0.554	0.063
p-value of (a)=(c)			0.007	0.545	0.142	0.000	0.002	0.003	0.190
p-value of (b)=(d)			0.000	0.000	0.000	0.593	0.001	0.032	0.002

Notes: This table reports results from Equation (1). The sample period is from 2008 to 2021. The estimated coefficients are multiplied by 1,000. The dependent variables are abnormal returns from 1 to 5 days after *Guba* opinions are published. $NPess_{it}$ and $NPos_{it}$ are proportions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day *t*. $TPess_{it}$ and $TPos_{it}$ are proportions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. $INews_{it}$ takes the value of 1 if firm *i* is covered by any newspapers on day *t* and 0 otherwise. $NegNews_{it}$, $PosNews_{it}$, and $Newsnum_{it}$ are the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, and the number of reading times and comments. All columns include the firm and year-by-month fixed effects. *P*-values in the bottom four rows are the *t*-statistic testing for the equality of the coefficients. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Effects of *Guba* Opinions on Overnight and Daytime Abnormal Returns

	Daily abnormal return		Overnight abnormal return				Daytime abnormal return	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$NPess_{it}$	−0.733***	−0.654***			−1.396***	−1.435***	0.690***	0.797***
	(−10.36)	(−8.99)			(−27.30)	(−26.57)	(9.64)	(10.69)
$NPos_{it}$	0.446***	0.400***			1.379***	1.232***	−0.255***	−0.242***
	(7.68)	(6.71)			(35.20)	(28.26)	(−4.52)	(−4.13)
$TPess_{it}$	−0.453***	−0.494***	−0.392***	−0.432***	−0.315***	−0.359***	−0.085	−0.109
	(−5.97)	(−6.16)	(−7.87)	(−8.00)	(−6.35)	(−6.66)	(−1.10)	(−1.35)
$TPos_{it}$	0.402***	0.411***	0.327***	0.242***	0.204***	0.140***	0.136*	0.128*
	(5.37)	(5.32)	(6.69)	(4.55)	(4.17)	(2.64)	(1.82)	(1.66)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,422,599	2,974,797	3,422,599	2,974,797	3,422,599	2,974,797	3,422,599	2,974,797
R^2	0.004	0.004	0.017	0.018	0.018	0.019	0.006	0.006

The sample period is from 2008 to 2021. The estimated coefficients are multiplied by 1,000. The dependent variables are daily returns (close-to-open, columns 1–2), overnight returns (close-to-open, columns 3–6), and daytime returns (open-to-close, columns 7–8). Columns 1–2 replicate the results from columns 3 and 5 in Table 3. $NPess_{it}$ and $NPos_{it}$ are proportions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day *t*. $TPess_{it}$ and $TPos_{it}$ are proportions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, and the number of reading times and comments. Newspaper controls include a dummy indicating that firm i is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. All columns include the firm and year-by-month fixed effects. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5. Effects of Online Opinions on Abnormal Returns (By Firm Size)

	Small: 1 (1)	2 (2)	3 (3)	4 (4)	Big: 5 (5)	*P*-value (6)
Panel A. Daily returns						
$NPess_{\mathrm{it}}$	−0.510**	−0.518***	−0.782***	−0.869***	−1.028***	0.000
	(−2.54)	(−3.16)	(−4.51)	(−4.97)	(−5.84)	
$NPos_{it}$	0.296*	0.161	0.237*	0.529***	0.894***	0.000
	(1.85)	(1.11)	(1.69)	(3.80)	(6.87)	
$TPess_{it}$	−0.584***	−0.127	−0.706***	−0.294	−0.528***	0.000
	(−2.63)	(−0.66)	(−3.83)	(−1.57)	(−3.00)	
$TPos_{it}$	0.473**	0.668***	0.269	0.402**	0.196	0.000
	(2.24)	(3.24)	(1.43)	(2.13)	(1.22)	
Controls	Yes	Yes	Yes	Yes	Yes	
Observations	544,339	453,771	539,143	626,039	811,505	
R^2	0.009	0.006	0.005	0.004	0.003	
Panel B. Overnight returns						
$NPess_{\mathrm{it}}$	−0.999***	−1.063***	−1.399***	−1.712***	−2.182***	0.000
	(−8.22)	(−8.94)	(−12.10)	(−14.05)	(−14.15)	
$NPos_{it}$	0.950***	1.015***	1.170***	1.395***	1.487***	0.000
	(9.05)	(10.20)	(12.06)	(13.98)	(16.63)	
$TPess_{it}$	−0.479***	−0.359***	−0.306**	−0.221*	−0.347***	0.000
	(−3.45)	(−2.69)	(−2.41)	(−1.90)	(−2.91)	
$TPos_{it}$	0.101	−0.066	0.165	0.102	0.087	0.017
	(0.73)	(−0.49)	(1.31)	(0.82)	(0.76)	
Controls	Yes	Yes	Yes	Yes	Yes	
Observations	544,339	453,771	539,143	626,039	811,505	
R^2	0.030	0.023	0.019	0.016	0.014	
Panel C. Daytime returns						
$NPess_{\mathrm{it}}$	0.564***	0.675***	0.520***	0.772***	1.238***	0.000
	(2.80)	(3.97)	(2.89)	(4.28)	(7.12)	
$NPos_{it}$	−0.08	−0.256*	−0.381***	−0.276**	−1.000	0.000
	(−0.51)	(−1.79)	(−2.66)	(−2.02)	(0.01)	
$TPess_{it}$	−0.104	0.216	−0.337*	0.005	−0.148	0.316
	(−0.48)	(1.10)	(−1.79)	(0.03)	(−0.84)	
$TPos_{it}$	0.217	0.563***	−0.062	0.146	−0.065	0.091
	(1.03)	(2.72)	(−0.33)	(0.77)	(−0.40)	
Controls	Yes	Yes	Yes	Yes	Yes	
Observations	544,339	453,771	539,143	626,039	811,505	
R^2	0.007	0.007	0.006	0.006	0.007	
Reversals (%)	56.4	63.5	37.2	45.0	56.7	

Notes: The sample period is from 2008 to 2021. The estimated coefficients are multiplied by 1,000. The dependent variables are daily returns (close-to-open, Panel A), overnight returns (close-to-open, Panel B), and daytime returns (open-to-close, Panel C). Columns 1–5 report the estimated coefficients on the four *Guba* opinion variables using samples with various firm sizes, which is calculated based on the prior June market capitalization. The last column reports the *p*-values of the F-test that check whether the coefficients are jointly equal to 0 among the previous columns are 0. $NPess_{it}$ and $NPos_{it}$ are fractions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day *t*. $TPess_{it}$ and $TPos_{it}$ are fractions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, the number of reading times and comments, a dummy indicating that the firm *t* is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. All columns include year-by-month and firm fixed effects. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6. Effects of Online Opinions on Earnings Surprise

	Earnings surprise: earnings per share		Earnings surprise: net profits	
	(1)	(2)	(3)	(4)
Panel A. Small firm				
$NPess_{it}$	−0.0378	0.0014	0.0002	−0.0201
	(−1.21)	(0.05)	(0.00)	(−0.77)
$NPos_{it}$	0.0341	0.0129	0.0607	0.0016
	(1.38)	(0.62)	(1.48)	(0.08)
$TPess_{it}$	−0.0844**	−0.0055	-0.0104	−0.0276
	(−2.57)	(−0.20)	(-0.19)	(−1.00)
$TPos_{it}$	−0.0405	−0.0051	0.0336	−0.0107
	(−1.22)	(−0.18)	(0.61)	(−0.39)
Controls	Yes	Yes	Yes	Yes
Observations	11,253	11,253	10,994	10,994
R^2	0.043	0.030	0.187	0.027
Panel B. Big firm				
$NPess_{it}$	−0.1182***	−0.0534*	-0.3494	−0.0404
	(−3.02)	(−1.77)	(-0.26)	(−1.34)
$NPos_{it}$	0.0788***	0.0654***	0.7631	0.0580***
	(2.71)	(2.92)	(0.76)	(2.58)
$TPess_{it}$	−0.0628	−0.0336	-0.7618	−0.0309
	(−1.52)	(−1.06)	(-0.54)	(−0.97)
$TPos_{it}$	0.1019***	0.0322	-0.1647	0.0442
	(2.59)	(1.06)	(-0.12)	(1.45)
Controls	Yes	Yes	Yes	Yes
Observations	10,295	10,295	10,297	10,297
R^2	0.032	0.021	0.058	0.111

Notes: The sample period is from 2008 to 2021. The dependent variables are earnings surprise, which is measured by the difference between the reported EPS and the average EPS forecasted by analysts (columns 1–2) and the difference between the reported profits and the average profits forecasted by analysts (columns 3–4). Columns 1 and 3 use continuous variables as the dependent variables. The earnings surprise of net profits is measured in 1 million RMB. Columns 2 and 4 use dummies as dependent variables, which take the value of 1 if the corresponding measure of earnings surprise is positive. Panels A and B report the estimated coefficients on the four *Guba* opinion variables using samples of small and big firms, which are categorized according to the median of the market capitalization on the prior June. $NPess_{it}$ and $NPos_{it}$ are proportions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day *t*. $TPess_{it}$ and $TPos_{it}$ are proportions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. Controls include the number of reading times and comments, a dummy indicating that firm i is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles. All columns include firm fixed effects. *P*-values in the bottom four rows are from the *t*-statistic testing for the equality of the coefficients. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7. Effects of *Guba* Opinions on Abnormal Returns (By Firms' Announcements of Major Events)

	$Ret_{i,t+1}$ (1)	$Ret_{i,t+2}$ (2)	$Ret_{i,t+3}$ (3)	$Ret_{i,t+4}$ (4)	$Ret_{i,t+5}$ (5)
Panel A. Controlling for the announcement of major events					
$NPess_{it}$	−0.653***	0.113	0.072	0.015	−0.087
	(−8.97)	(1.48)	(0.92)	(0.19)	(−1.10)
$NPos_{it}$	0.978***	0.333***	0.193***	0.187***	0.178***
	(16.44)	(5.30)	(3.04)	(3.01)	(2.86)
$TPess_{it}$	−0.496***	−0.518***	−0.299***	−0.331***	−0.245***
	(−6.19)	(−5.92)	(−3.39)	(−3.87)	(−2.82)
$TPos_{it}$	0.412***	0.391***	0.554***	0.416***	0.515***
	(5.34)	(4.57)	(6.59)	(4.85)	(5.87)
Major event	−0.191***	0.021	−0.052	0.025	−0.257***
	(−3.61)	(0.39)	(−0.99)	(0.50)	(−4.96)
	Yes	Yes	Yes	Yes	Yes
Observations	2,974,797	2,743,955	2,771,363	2,763,199	2,756,885
R^2	0.004	0.002	0.003	0.003	0.003
Panel B. Opinions published on major-event days					
(a) $NPess_{it}$	−0.710***	−0.048	0.043	−0.025	−0.097
	(−3.49)	(−0.16)	(0.14)	(−1.10)	(−0.96)
(b) $NPos_{it}$	1.162***	0.652***	0.458**	0.196**	0.035
	(4.36)	(2.85)	(2.07)	(−2.01)	(0.16)
(c) $TPess_{it}$	−0.760***	−0.937***	−0.638***	−0.637*	0.111
	(−4.08)	(−2.75)	(−3.40)	(−1.91)	(0.32)
(d) $TPos_{it}$	0.330	0.391	0.348	0.240	0.597*
	(1.10)	(1.20)	(1.04)	(0.74)	(1.87)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	257,513	237,060	238,775	244,330	237,812
Adjusted R^2	0.011	0.003	0.003	0.004	0.004

Panel C. Opinions published on regular days					
(e) $NPess_{it}$	−0.391***	0.116	0.068	0.038	−0.077
	(−9.12)	(1.48)	(0.83)	(0.46)	(−0.94)
(f) $NPos_{it}$	0.352***	0.238***	0.113*	0.164**	0.133**
	(5.75)	(3.65)	(1.73)	(2.51)	(2.04)
(g) $TPess_{it}$	−0.427***	−0.493***	−0.231**	−0.310***	−0.272***
	(−5.67)	(−5.48)	(−2.53)	(−3.50)	(−3.01)
(h) $TPos_{it}$	0.451***	0.364***	0.552***	0.408***	0.496***
	(5.61)	(4.12)	(6.31)	(4.59)	(5.42)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,717,284	2,506,895	2,532,588	2,518,869	2,519,073
R^2	0.004	0.002	0.003	0.003	0.003
p-value of (a)=(e)	0.000	0.154	0.386	0.899	0.270
p-value of (b)=(f)	0.000	0.000	0.026	0.327	0.036
p-value of (c)=(g)	0.000	0.000	0.000	0.000	0.006
p-value of (d)=(h)	0.120	0.873	0.131	0.147	0.633

Notes: The sample period is from 2008 to 2021. The estimated coefficients are multiplied by 1,000. In Panel A, we control for a dummy indicating for firms' announcement of major events. In Panel B, we estimate the sample of opinions published on days when firms announce major events to test their predictability on return on the next day. In Panel C, we estimate the sample of opinions when firms do not publish major events. The dependent variables are abnormal returns from 1 to 5 days after *Guba* opinions are published. $NPess_{it}$ and $NPos_{it}$ are proportions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day *t*. $TPess_{it}$ and $TPos_{it}$ are proportions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. Controls include the number of reading times and comments, a dummy indicating that firm *i* is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles. All columns include the firm and year-by-month fixed effects. P-values in the bottom four rows are from the t-statistic testing for the equality of the coefficients between those in Panels B and C. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8. Effects of Online Opinions on Investor Attention (Interacted with Major Events)

	Reading times				Number of comments			
	Non-trading-hour reads	Trading-hour reads	Overall reads	Overall reads	Non-trading-hour comments	Trading-hour comments	Overall comments	Overall comments
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of non-trading-hour articles	1,089.665***		1,216.892***	1,146.444***	2.533***		2.876***	2.753***
	(7.27)		(6.48)	(6.47)	(5.66)		(5.74)	(5.57)
# of trading-hour articles		492.323***	620.318***	634.943***		1.155***	1.449***	1.470***
		(11.13)	(8.77)	(9.14)		(20.19)	(14.19)	(13.72)
# of non-trading-hour articles×major event				602.895***				1.112***
				(3.42)				(4.26)
# of trading-hour articles×major event				−206.453***				−0.352***
				(−4.33)				(−3.92)
Major event				7,266.445**				−4.62
				(1.96)				(−0.71)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,974,797	2,974,797	2,974,797	2,974,797	2,974,797	2,974,797	2,974,797	2,974,797
R^2	0.034	0.31	0.082	0.083	0.324	0.033	0.136	0.137

Notes: The sample period is from 2008 to 2021. The dependent variable is the number of reading times (columns 1–4) and the number of comments (columns 5–8). "# of non-trading-hour articles" and "# of trading-hour articles" are the number of stock-*i* related non-trading-hour pessimistic and positive articles published on day *t*. "Major event" is a dummy that takes the value of 1 if a firm releases major events on day *t*. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, a dummy indicating that firm *i* is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. All columns include the firm and year-by-month fixed effects. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9. Effects of Online Opinions on Investor Attention (Interacted with Volatility)

	Reading times		Comments	
	(1)	(2)	(3)	(4)
# of non-trading-hour articles	1,211.237***	534.584***	2.866***	1.524***
	[6.38]	[5.91]	[5.67]	[4.66]
# of trading-hour articles	606.245***	754.466***	1.419***	1.653***
	[8.49]	[8.74]	[14.26]	[6.38]
# of non-trading-hour articles×volatility		18,977.053***		37.549***
		[4.17]		[5.92]
# of trading-hour articles×volatility		−5,477.695***		−9.345***
		[−4.23]		[−2.70]
volatility		71,605.07		−185.84
		[0.47]		[−0.66]
Controls	Yes	Yes	Yes	Yes
Observations	2,974,797	2,974,797	2,974,797	2,974,797
R^2	0.083	0.085	0.137	0.14

Notes: The sample period is from 2008 to 2021. The dependent variable is the number of reading times (columns 1–2) and the number of comments (columns 3–4). "# of non-trading-hour articles" and "# of trading-hour articles" are the number of stock-*i* related non-trading-hour pessimistic and positive articles published on day *t*. "volatility" is stock *i*'s volatility, which is measured by the past five-day rolling window standard deviation of stock returns. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, a dummy indicating that firm *i* is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. All columns include the firm and year-by-month fixed effects. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10. Robustness Check

	$Ret_{i,t+1}$ (1)	$Ret_{i,t+2}$ (2)	$Ret_{i,t+3}$ (3)	$Ret_{i,t+4}$ (4)	$Ret_{i,t+5}$ (5)
Panel A. Orthogonal opinion measures					
$NPess_{it}$	−0.783***	0.008	0.077	−0.036	−0.139
	(−9.03)	(0.09)	(0.87)	(−0.41)	(−1.59)
$NPos_{it}$	1.102***	0.196***	0.134*	0.100	0.050
	(16.17)	(2.82)	(1.87)	(1.45)	(0.72)
$TPess_{it}$	−0.441***	−0.481***	−0.308***	−0.250**	−0.264***
	(−4.93)	(−5.01)	(−3.13)	(−2.58)	(−2.71)
$TPos_{it}$	0.411***	0.378***	0.421***	0.432***	0.453***
	(4.62)	(3.92)	(4.50)	(4.47)	(4.68)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,801,093	2,640,123	2,634,104	2,628,561	2,641,895
R^2	0.004	0.002	0.003	0.003	0.002
Panel B. Controlling for firm characteristics					
$NPess_{it}$	−0.692***	0.065	0.03	−0.028	−0.123
	(−9.49)	(0.85)	(0.38)	(−0.36)	(−1.56)
$NPos_{it}$	0.946***	0.315***	0.185***	0.176***	0.167***
	(15.85)	(5.02)	(2.92)	(2.85)	(2.68)
$TPess_{it}$	−0.468***	−0.507***	−0.291***	−0.321***	−0.227***
	(−5.83)	(−5.78)	(−3.28)	(−3.74)	(−2.62)
$TPos_{it}$	0.401***	0.383***	0.551***	0.415***	0.515***
	(5.18)	(4.47)	(6.54)	(4.84)	(5.87)
log ($size_{it}$)	−1.227***	−1.309***	−1.286***	−1.192***	−1.105***
	(−17.99)	(−19.31)	(−19.31)	(−18.54)	(−17.30)
B/M_{it}	0.083	0.287**	0.587***	0.563***	0.717***
	(0.70)	(2.26)	(4.83)	(4.63)	(5.88)
$Trading\ volumn_{it}$	−0.136***	−0.058***	−0.017**	−0.026***	−0.037***
	(−8.14)	(−5.04)	(−1.98)	(−2.70)	(−4.10)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,974,797	2,743,955	2,771,363	2,763,199	2,756,885
R^2	0.005	0.003	0.003	0.003	0.003

Panel C. Controlling for industry-by-year fixed effects

$NPess_{it}$	−0.646***	0.119	0.076	0.016	−0.083
	(−8.87)	(1.56)	(0.96)	(0.20)	(−1.05)
$NPos_{it}$	0.986***	0.339***	0.198***	0.193***	0.189***
	(16.57)	(5.39)	(3.12)	(3.11)	(3.03)
$TPess_{it}$	−0.485***	−0.511***	−0.292***	−0.327***	−0.236***
	(−6.05)	(−5.83)	(−3.31)	(−3.83)	(−2.70)
$TPos_{it}$	0.420***	0.398***	0.561***	0.422***	0.525***
	(5.44)	(4.65)	(6.67)	(4.92)	(5.99)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,974,797	2,743,955	2,771,363	2,763,199	2,756,885
R^2	0.005	0.003	0.003	0.003	0.003

Notes: This table reports results from Equation (1). The sample period is from 2008 to 2021. The estimated coefficients are multiplied by 1,000. Panel A uses a two-step method to control the autocorrelation of opinions as well as the correlation between stock return on date *t* and *Guba* opinions on date *t*. In the first step, we regress each of our opinion variables on stock returns from *t*–5 to *t*, opinion variables from *t*–5 to *t*, controls, firm and month-by-year fixed effects. In the second step, we replace the opinion variables with the residuals obtained from the first step in Equation (1). Panel B includes additional controls for market value in logarithm, book-to-market ratio, and trading volume. Trading volume is measured in 100 million RMB. $NPess_{it}$ and $NPos_{it}$ are proportions of stock-*i* related non-trading-hour pessimistic and positive opinions published on day t. $TPess_{it}$ and $TPos_{it}$ are proportions of stock-*i* related trading-hour pessimistic and positive opinions published on day *t*. Controls include abnormal returns from two days before the opinions are published to the day when opinions are published, the last three-month holding period abnormal returns, the number of reading times and comments, a dummy indicating that firm *i* is covered by newspapers, and the number of firm-*i* related pessimistic and positive articles, and the total number of newspaper articles published on day *t*. All columns include the firm and year-by-month fixed effects. Standard errors are clustered at the firm level and *t* statistics are in the parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

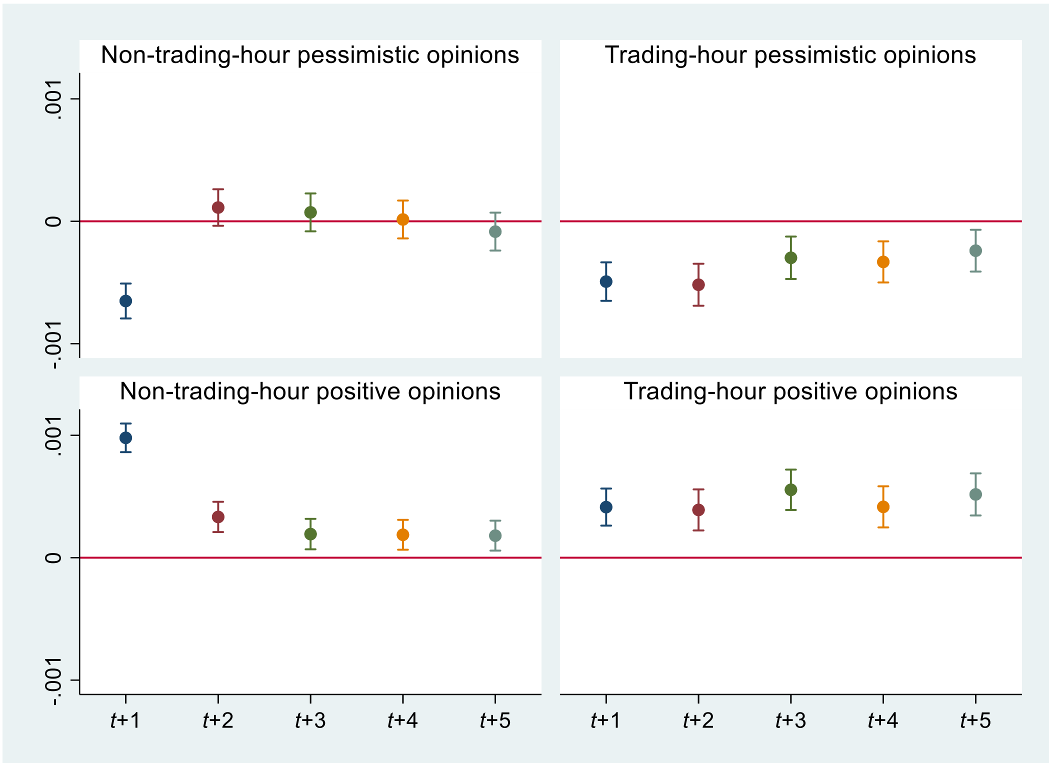

Figure 1. Estimated Coefficients of Online Opinions on Abnormal Returns in the Short Run

Notes: This figure illustrates the estimated coefficients and 95% confidence intervals by regressing abnormal returns from 1 to 5 days after the *Guba* opinions are published on the proportion of trading- and non-trading-hour pessimistic and positive opinions.

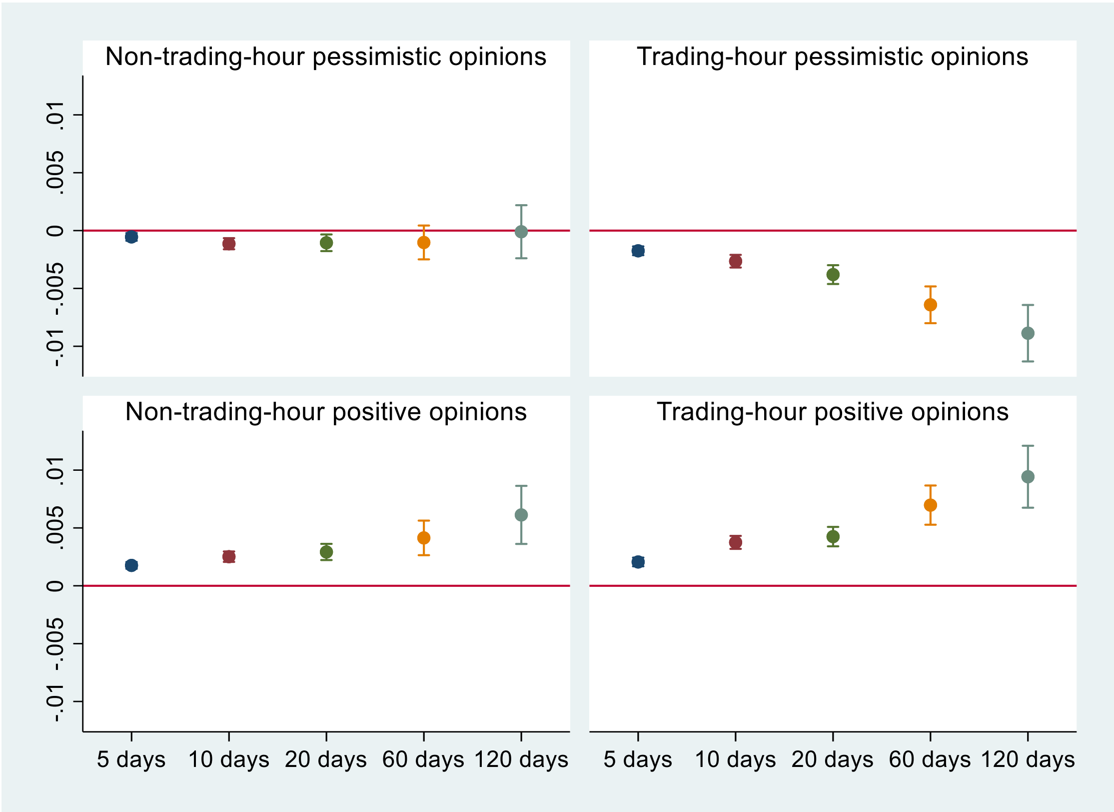

Figure 2. Estimated Coefficients of Online Opinions on Cumulative Abnormal Returns in the Long Run

Notes: This figure illustrates the estimated coefficients and 95% confidence intervals by regressing CARs over the 5-, 10-, 20-, 60-, and 120-day windows after *Guba* opinions are published on the proportion of trading- and non-trading-hour pessimistic and positive opinions.

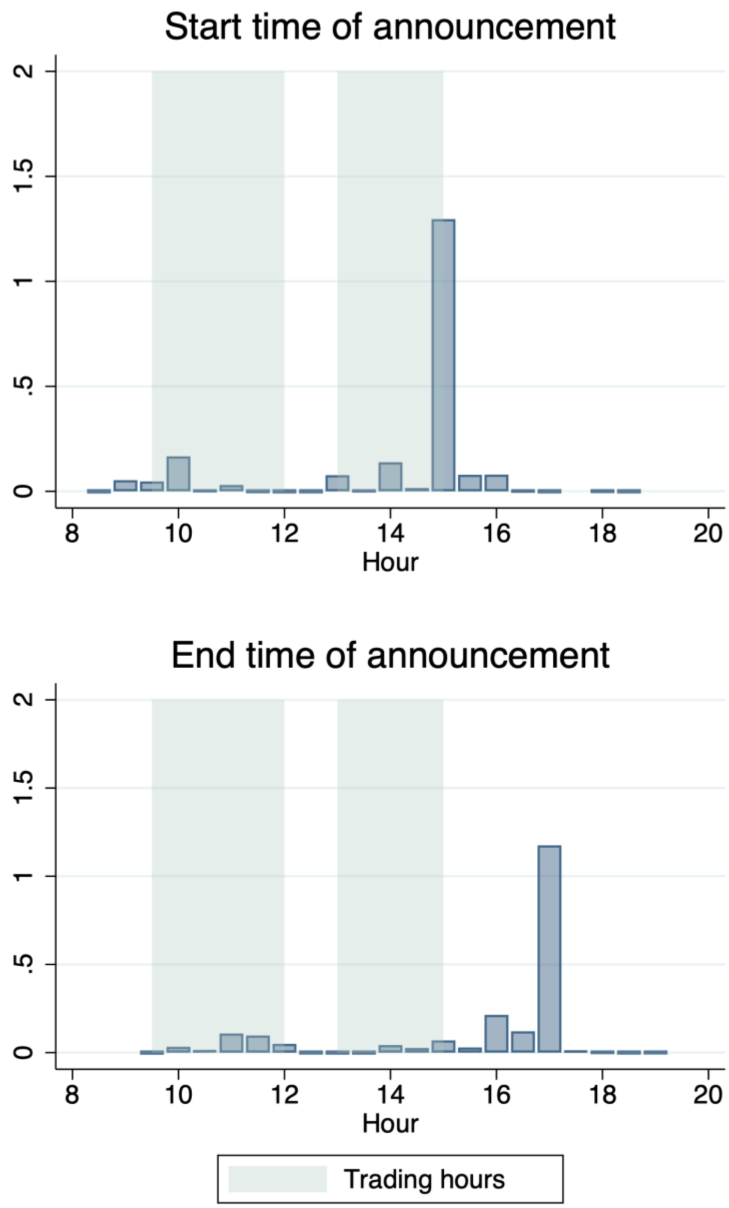

Figure 3. Distribution of Start and End Times of Press Conferences

Notes: The figure illustrates the distribution of hours of press conferences where companies announce major events. The data are collected from 2012 to 2021.

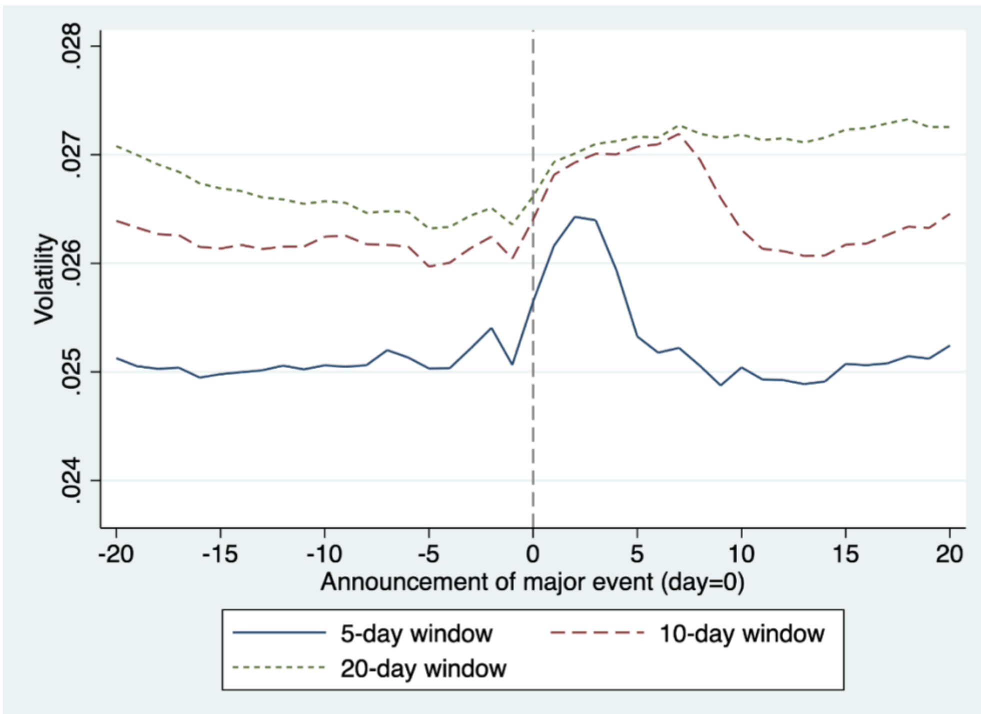

Figure 4. Event Study of Volatility Around the Major-event Announcement Day

Notes: The figure illustrates the volatility of stock around firms' announcement of major events. Day 0 represents firms' announcement of major events. The volatility on day *t* is calculated by taking the standard deviation of stocks in the window from *t*–4 to *t* (the blue line), *t*–9 to *t* (the dashed line), and *t*–19 to *t* (the dot line). We then roll the window to calculate the volatility 20 days before and after the announcement of a major event.

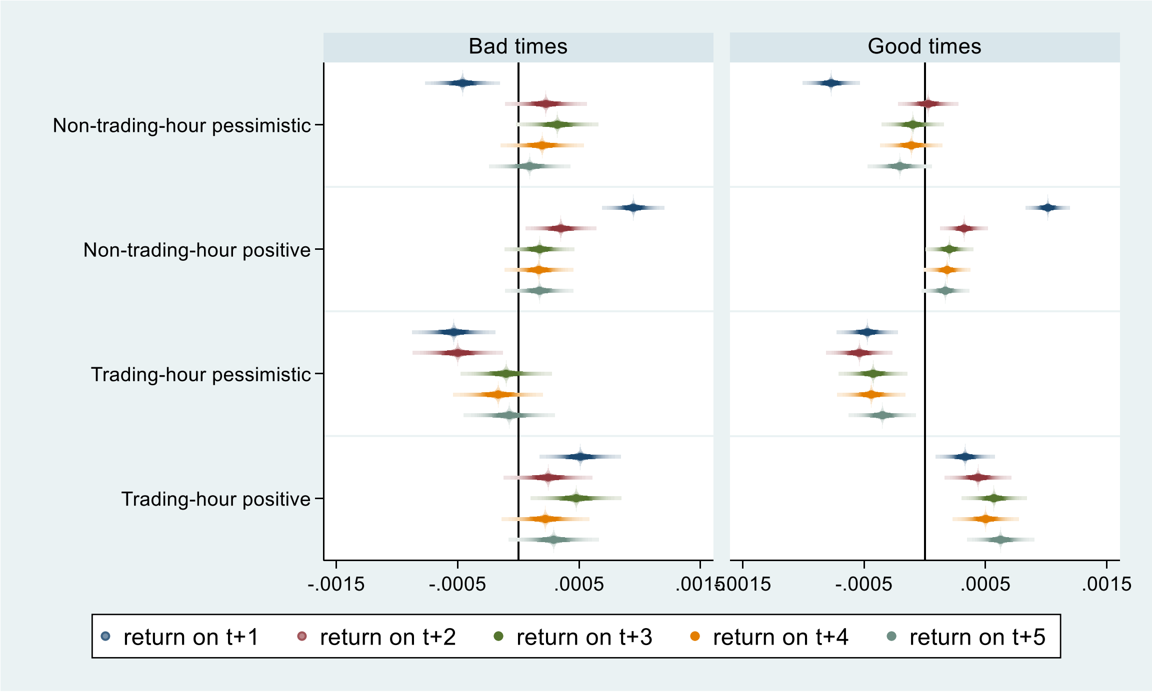

Figure 5. Estimated Coefficients in Good and Bad Times

Notes: The figure reports the estimated coefficients and 95% confidential intervals of our four *Guba* opinion variables using Equation (1). The outcome variables are abnormal returns from *t*+1 to *t*+5. The results are separately estimated in good and bad times.