Journal of Economic Behavior and Organization 197 (2022) 493–518

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Positive information shocks, investor behavior and stock price crash risk

Xin Cui[a], Ahmet Sensoy[b], Duc Khuong Nguyen[c,d], Shouyu Yao[a,*], Yiyao Wu[a]

[a] College of Management and Economics, Tianjin University, Tianjin, China
[b] Faculty of Business Administration, Bilkent University, Ankara, Turkey
[c] IPAG Business School, Paris, France
[d] International School, Vietnam National University, Hanoi, Viet Nam

ARTICLE INFO

Article history:
Received 5 June 2021
Revised 19 January 2022
Accepted 15 March 2022
Available online 31 March 2022

JEL classifications:
G10
G18

Keywords:
Positive information shock
Crash risk
Retail investors

ABSTRACT

This article explores the impact of positive information shocks on investors' trading behavior and the related stock price crash risk. We use cumulative positive jump returns to measure the positive information shocks and find that these shocks exacerbate crash risk. Moreover, retail investor attention, over-optimistic investor sentiment, and retail trades are channels for this exacerbation. We also provide evidence that the effect of the information shocks varies across firm characteristics and aggregate states. It is stronger for firms with large-cap, long listing times, and state-owned structures and during over-optimistic aggregate states. Overall, our results shed light on investor trading behavior and market risk related to unexpected information shocks, which helps detect and diagnose potential market instability.

1. Introduction

The market risk caused by the extreme stock price jump is a hot issue, especially in a financial crisis period. In a financial crisis, the stock price will become extremely sensitive, especially for some unexpected information shocks, which can cause a huge jump in the price and lead to more severe market consequences (Jiang and Kim, 2016). For example, DeBondt and Thaler (1985) suggest that most investors tend to overreact to unexpected news, resulting in the long-term reversals of stock returns Savor (2012). finds that price shocks strongly correlate with aggregate implied volatility. In addition, Jarociński and Karadi (2020) point out that central bank announcements can convey an assessment of the economic outlook and find that the surprises in these assessments, namely "central bank information shocks," have large impacts on stock prices and the economic stability.

However, previous studies typically examine all news shocks together and thus get a sort of average of quite different effects (Chan, 2003; Savor, 2012; Tetlock, 2010, 2011). Only a few studies have explored the market reaction to adverse information shocks (Park and Lee, 2014; Niu and Zhang, 2021). Moreover, relevant studies rarely investigate investor behavior and market risk accompanied by information shocks. Compared with negative news, the impact of positive information is released and absorbed quickly, and the market reaction is stronger (Park and Lee, 2014;), which may distort investor

* Corresponding author.
E-mail address: yaosy@tju.edu.cn (S. Yao).

https://doi.org/10.1016/j.jebo.2022.03.016
0167-2681/

behavior and cause stronger market risk. In this article, we extend this literature and use unexpected and dramatic positive jumps in stock prices as the measure of positive information shocks to explore the effects of positive information shocks on investor reactions and market risk.

Our analysis is motivated by two opposing strands of research concerning the effect of positive information shocks on market reactions. In the first strand, Frank and Sanati (2018) point out that retail investors tend to overreact to positive news, thus implying that these shocks are followed by reversal. Accordingly, these investors would pay more attention to stocks with frequent positive price jumps, be highly optimistic and be the net buyers of these stocks (Yao et al., 2019; Cheng et al., 2021), pushing the prices up to higher levels. However, with the continuous integration of information, once retail investors realize that stock prices have been overestimated by the market, they would sell those stocks off, causing price bubbles to burst, which may further exacerbate stock price crash risk (Yin and Tian, 2017).

Some studies in the second strand of the literature suggest that positive information conveys the signal to outside investors that a firm's fundamentals have been improving (e.g., Skinner and Sloan, 2002). This improvement is typically conductive to the sustained growth of stock returns (Savor, 2012; Novy-Marx, 2013; Asness et al., 2019), which reduce stock price crashes (Chen et al., 2017). In addition, positive information shocks can attract more investor attention. As a tool for external monitoring, investor attention helps reduce managers' bad behavior and mitigate information asymmetry (Brown et al., 2009), thus lowering the crash risk (Wen et al., 2019; Wang et al., 2020).

It can be seen from the above analysis that the views regarding the potential consequences of positive information shocks are mixed. It is still an important to empirically investigate whether and how positive information shocks affects investor trading behavior and related market risk. Using data from the Chinese stock market, our study conducts an in-depth analysis of these issues and allows to better understand the underlying market dynamics following the positive information shocks. Meanwhile, the in-depth analysis of investor behavior in this paper is also conducive to the corresponding market risk prevention, especially in the period of major financial crisis when prices are more sensitive to information.

Several reasons motivate our investigation in the context of the Chinese market. First and foremost, the characteristics of Chinese stock market provides a good setting to examine information shocks and crash risk. Indeed, although the Chinese stock market started late, it develops rapidly and has surpassed the European Union and Japan to become the world's second largest stock market after the United States (Brzeszczyński et al., 2015; Yao et al., 2019) in terms of market capitalization. However, China's current institutional background, regulatory system, and investor literacy cannot keep up with the rapid development of the capital markets. With poor stability and self-correction ability, it is vulnerable to the impact of information shocks, and it experiences sharp rises and falls in stock prices frequently. Moreover, different from developed markets, the Chinese stock market is still an emerging market with partial integration and dominant ownership by local investors (Nartea et al., 2017; Yao et al., 2019). This special environment of the Chinese market thus allows us to observe the trading behavior of a completely different set of investors (Nartea et al., 2017). In addition, unlike fully developed stock markets, the Chinese stock market is dominated by retail investors which lack sophisticated information processing and expertise in investment analysis (Gao and Yang, 2018), and often show speculative characteristics such as gambling preference (Lin and Liu, 2017; Yao et al., 2019), and abnormal optimism and excessive trading (Han and Li, 2017; Li et al., 2017). Accordingly, our study can uncover the retail trading behavior and the evolution of market risk in response to information shocks. Strict short sale constraints are also an important reason for studying the Chinese context because the overpricing caused by divergent opinions cannot be quickly corrected (Cheema and Nartea, 2014; Chang et al., 2014). When overvalued prices return to fundamental values and the accumulated price bubble bursts, the crash risk might increase. Recently, the Chinese market has relaxed short-selling restrictions. This policy shock is similar to a natural experiment for further exploring the causal relationship between retail trading behavior and market risk.

Our analysis shows that positive information shocks significantly aggravate the future crash risk whereby the short-term effect is stronger than the long-term effect. Moreover, this finding remains intact in a series of robustness and endogeneity tests such as the GMM model, the instrumental variable, and the time-varying fixed-effect model controlling for potential omitted variables. By further analyzing the role of retail investors, we find that retail investor attention attracted by positive information shocks does not play the monitoring role in reducing information asymmetry. By contrast, positive information shocks retain the attention of retail investors, stimulating their excessive optimism sentiment and aggressive trading behavior, and aggravating the stock price crash risk in the future. Our results are similar to those of Frank and Sanati (2018) who find that retail investors overreact to positive news, and a positive news is followed by a reversal of stock return.

In order to further verify the role of retail investors, we also use exogenous policy shocks to explore the causality. We find that exogenous policies, such as relaxing short sale constraints and implementing internet speech restrictions, can effectively curb retail speculation and weaken the exacerbating effect of positive information shocks on the crash risk. Besides, the effect of positive information shocks varies across firm characteristics and aggregate states. For example, it is more pronounced for state-owned enterprises and firms with high market capitalization and a long listing time. In addition, taking external market environment into account, we find that the effect of positive information shocks on crash risk is more likely to be strengthened under the condition of market development, high market sentiment, and bullish market trend.

Overall, our study's main contributions are threefold. We first contribute to the literature on information shocks and stock price reaction (Chan, 2003; Tetlock, 2010, 2011; Savor, 2012; Jiang and Zhu, 2017; Frank and Sanati, 2018), and especially positive information shocks. As we mentioned earlier, previous studies have investigated either all news shocks together (Chan, 2003; Savor, 2012; Tetlock, 2010, 2011) or market reaction to negative information shocks (Park and Lee, 2014;

Brennan et al., 2016; Niu and Zhang, 2021). Most studies based on the U.S. market show evidence of market underreaction to information shocks (Chan, 2003; Savor, 2012; Govindaraj et al., 2013; Jiang and Zhu, 2017), while some few others document market overreaction to information shocks (Tetlock, 2010, 2011; Engelberg et al., 2012). Our study thus extends these studies by showing that investors significantly overreact to positive information shocks and that considering the sign of information shocks help reconcile the previous seemingly opposite findings (Frank and Sanati, 2018).

Second, our study adds to the strand of literature focusing on the trading mechanisms behind the price formation process. Previous studies only provide potential explanations for market reaction after information shocks, without providing direct evidence (Chan, 2003; Savor, 2012; Jiang and Zhu, 2017). Differently, we explore in detail the retail investor behavior in the face of positive information shocks and provide direct evidence about changes in investor sentiment and trading behavior, thus complementing research on investor sentiment and financial market behavior (Barberis et al., 1998; Ftiti et al., 2016; Jawadi et al., 2018). More importantly, we show that considering the heterogeneity of investors rather than market aggregation provides a more comprehensive understanding of various market shock effects.

Finally, our research explores the changes in stock price crash risk with respect to information shocks and thus extends the literature on possible causes of crash risk. Prior literature examines the crash risk determinants from the viewpoints of "bad news holding" and "investor behavior" (Jin and Myers, 2006; Kaplanski and Levy, 2010). Our study proposes competitive hypotheses and focuses on the effects on crash risk of positive information shocks, measured by large positive discontinuous changes in stock prices. Furthermore, whereas previous studies use annual and quarterly data to measure crash risk (Cheng et al., 2020), we improve the traditional calculation method and obtain the monthly measure of crash risk. This improvement allows us to dynamically grasp the short-term changes of crash risk.

The remainder of the article is organized as follows Section 2. proposes literature review and contradictory hypotheses Section 3. presents the data and empirical model Section 4. analyzes the impact of positive information shocks on crash risk and presents robustness results Section 5. examines the economic channels between positive information shocks and crash risk Section 6. performs further analysis Section 7. concludes the article and provides some policy implications.

2. Literature review and hypotheses development

2.1. Information shocks and stock market reaction

Prior literature has examined the relation between information shocks and stock returns (Savor, 2012; Jiang and Yao, 2013), price discovery (Jiang et al., 2011), analyst recommendation revisions (Conrad et al., 2006; Jiang and Kim, 2016), and short-term market reaction (Jiang and Zhu, 2017). While the majority of studies consider all information shocks together (Chan, 2003; Tetlock, 2011; Savor, 2012), only some of them pay attention to positive and negative information shocks separately (Park and Lee, 2014; Jiang and Zhu, 2017; Ma et al., 2021). Since the direction of information shocks is not distinguished in most earlier studies, the literature has formed two completely opposite conclusions about stock price reaction following information shocks: underreaction (Chan, 2003; Savor, 2012; Govindaraj et al., 2013; Jiang and Zhu, 2017) and overreaction (Tetlock, 2010, 2011; Engelberg et al., 2012). For example, Savor (2012) and Govindaraj et al. (2013) find underreaction to information shocks that leads to drifts in returns over the next several months. On the contrary, Tetlock (2010) finds reversals after news shocks when considering 10-day post-shock return patterns and interprets this finding as overreaction to new information.

Recent research argued that market participants have different responses to positive and negative information shocks (Park and Lee, 2014; Ma et al., 2021), which may cause different market reactions. As emphasized by Gao et al. (2018), bad news travels slowly as it takes longer time for the market to digest the informational content of pessimistic news Frank and Sanati (2018). also show that negative information is followed by drift. Compared with the expected decline of stock prices brought by negative information shocks (Hong et al., 2000; Gao and Yang, 2018), a series of chain market reactions brought by positive information shocks is more difficult to predict (Jiang and Zhu, 2017; Frank and Sanati, 2018). The impact of positive information is also released and absorbed quickly, and the market reaction is stronger (Park and Lee, 2014; Brennan et al., 2015), which may distort investor behavior and cause stronger market risk. Therefore, financial risks and potential crises followed by positive news are more difficult to manage.

2.2. Stock price crash risk and theoretical background

When it comes to stock price crashes, it not only harms stockholders' wealth, but also has a great shock on the stock market and the real economy. Exploring the factors influencing the crash risk is, thus, of paramount importance. On the one hand, earlier studies have examine how bad news hoarding affect stock price crash risk and point out that due to information asymmetry, firm managers tend to hide bad news for their career development and self-interest (Kim et al., 2016; Jia, 2018). When bad news is too important to hide or bad news is revealed, stock prices inevitably suffer from downward pressure, resulting in price crashes (Chen et al., 2018). Based on the principal-agent theory, the existing literature also explores the determinants of crash risk from tax avoidance (Kim et al., 2011a), CSR (Kim and Kim, 2014), liability insurance (Yuan et al., 2016), CEO overconfidence (Kim et al., 2016), and debt financing (Wang et al., 2020), to name a few.

On the other hand, there are studies investigating the influencing factors of crash risk according to the “retail investor behavior”. Retail investors are sensitive to market sentiment, and they contribute to drive up stock prices when they are highly optimistic, causing stock price bubbles (Fu et al., 2021; Chen and Schmidt, 2021). Then, when the bubbles burst, stock prices decline sharply and crash risk may happen. Previous studies have found that heterogeneous investor beliefs (Cao and Ouyang, 2005), herding effects (Xu et al., 2017; Deng et al., 2018), and investor sentiment (Kaplanski and Levy, 2010; Jang and Kang, 2019) are significantly associated with crash risk.

2.3. Hypothesis development

From the perspective of the “bad news hoarding” argument, positive information shocks may reduce stock price crash risk. Positive price changes can be driven by the improvement of firms’ fundamentals and prospects (Savor 2012). Improved fundamentals reflect better firms’ operation and lower likelihood of concealing bad news, causing a lower crash risk (Chen et al., 2017).

Moreover, large positive stock price jumps are attention-grabbing events and can catch more investor attention (Jiang and Zhu, 2017). The crash risk is expected to reduce since this external monitoring helps alleviate information asymmetry between investors and managers Barber and Odean (2007). indicate that due to their limited cognitive ability, investors only analyze the information of stocks that attract their attention and those that are in their selection range. Investor attention may be distracted prior to the arrival of positive information shocks, but once these shocks arrive, attention would be focused on individual stocks. When investors start to pay attention to a stock, they can effectively play the external monitoring role Chen et al. (2007). argue that investor monitoring helps gather firm-specific information. Based on the “bad news hoarding” theory, we further argue that the gathering of a firm’s information by investors leads to diminish bad news hoarding and is likely to decrease the crash risk (Shleifer and Vishny, 1986; Wang and Zhang, 2009; Ding and Hou, 2015; Wen et al., 2019; Wang et al., 2020).

In sum, stocks with positive information shocks not only represent the improvement of firms’ fundamentals, but also attract more investor attention which can play the monitoring role in mitigating information asymmetry and lowering crash risk. This argument leads us to formulate the first hypothesis as follows:

H1: Positive information shocks reduce future crash risk.

As mentioned earlier in the introduction, the second strand of the literature suggests that positive information shocks may exacerbate crash risk due to the irrational behavior of retail investors. In this regard, Vozlyublennaia (2014) argues that a shock to returns may provoke changes in investor attention. Similarly, Yao et al. (2021) indicate that the maximum daily stock return of a company will greatly attract the attention of individual investors. Moreover, as for the market consequences of abnormal attention of individual investors, Da et al. (2011) point out that more retail investor attention can create extra noise to stock prices. Therefore, after a positive jump, the continued rise of stock prices may not be driven by firms’ positive fundamentals, but the positive information shocks that have attracted retail investor attention (Nofsinger, 2001).

Furthermore, Cheng et al. (2021) show that retail investor attention is closely related to their sentiment and trading, and that more positive investor attention is often accompanied by the increase of investor sentiment. Many studies have identified investor sentiment as an important factor affecting investor trading behavior and asset prices (Shleifer and Summers, 1990; Campbell and Kyle, 1993; Lakonishok et al., 1994; Barberis et al., 1998; Jawadi et al., 2018). These studies find that retail traders are likely to affect stock prices and investor’s decision-making process based on unpredictable changes in investor sentiment (Loewenstein, 2001; Böhm and Chiarella, 2005; Ftiti et al., 2016). In particular, De Long et al. (1990) indicate that when noise traders have high sentiment for a stock, they may push the price to a higher level. In addition, Barber et al. (2008) argue that changes in investor sentiment will lead to excessive transactions of retail investors and lead to mispricing.

It can be inferred from the above literature that abnormal retail attention, abnormal sentiment, and excessive trading behavior drive stock prices to rise continuously (Kim and Kim, 2014; Lin and Liu, 2017). Moreover, the aggressive individual investor behavior will limit arbitrage opportunities, hinder timely price correction, and further trigger corresponding speculation (Yao et al., 2019). However, with the consolidation of information, stock prices no longer have upward support and face severe crash risk when the abnormal attention, over-optimistic sentiment, and excessive trades cease (Yin and Tian, 2017). Some earlier studies provide preliminary evidence to support this argument Frank and Sanati (2018). find, for example, that retail investors overreact to positive news, but positive price shocks are followed by reversal Lin and Liu (2017). document that retail investors are attracted by high return stocks for speculative trading, but future returns on those stocks are significantly lower Yin and Tian (2017)., and Jiang and Kang (2019) find evidence of positive effect of sentiment-driven overpricing on future price crashes.

Overall, when retail investors are attracted by positive information shocks, abnormal retail attention, over-optimistic sentiment, and excessive trades lead to overvalued stock prices. This speculative retail investor behavior bias aggravates crash risk. Therefore, we propose to test the second hypothesis as follows:

H2: Positive information shocks exacerbate future crash risk.

3. Data and methodology

3.1. Data

Our dataset covers all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018.[1] The data frequency used in this study includes monthly, semi-annual, and annual. Stocks listed for less than one month during the study period are excluded. The stock trading data used for jump test and identification is from Wind database,[2] while the data used for calculating crash risk indicators and related control variables is mainly from the Chinese Stock Market and Accounting Research (CSMAR)[3] database and the RESSET[4] financial database.

3.2. Identification: the measure of positive information shocks

Previous literature commonly uses corporate events, analyst reports, and news stories to measure information shocks (Conrad et al., 2006; Savor, 2012; Frank and Sanati, 2018). More recent research uses large price changes as proxies for information events (e.g., Conrad et al., 2006; Savor, 2012) Jiang and Oomen (2008). put forward a new jump test approach and use the detected jumps in stock prices as a proxy for information shocks. Their jump test method stipulates that in the absence of jumps, the difference between simple and log returns equals one half of the instantaneous return variance. Previous studies such as Andersen et al. (2010), Jiang and Yao and (2013), Jiang and Zhu (2017), and Jiang et al. (2018) have adopted the same jump test method to carry out jump detection and measure the corresponding information impact.

Our study follows the Jiang and Oomen (2008)'s method to measure large positive discontinuous changes in stock prices (i.e., positive jumps) and use them as a proxy for positive information shocks. This method is advantageous in that it is model-free and not affected by the market microstructure noise (Jiang and Kim, 2016). It also differs from the Frank and Sanati (2018)'s approach because the event dates and the actual time of information arrival are not required. Moreover, this method is not restricted to public events and rather focuses on unexpected jumps that can cause large and discontinuous fluctuations in stock prices. Finally, the use of jumps to proxy significant information shocks captures market reaction to information shocks, instead of delayed reaction to corporate events (Jiang and Zhu, 2017) Jiang and Yao (2013). find that jumps provide a good measure of information shocks. Other studies such as Kapadia and Zekhnini (2019), Xiao et al. (2020), and DeLisle et al. (2021) have used jumps in stock prices as a proxy for large information shocks.

3.3. Cumulative positive jump returns

A price jump is a rare large stock price change caused by unexpected information shocks. Following Conrad et al. (2006), Savor (2012), and Jiang and Zhu (2017), we use cumulative positive jump returns (CJR) as the proxy for positive information shocks.

Based on the "variance swap" method of Jiang and Oomen (2008), we use the daily returns to build the jump test statistic JS_0 (for specific construction steps, please refer to Appendix B). The null hypothesis of the jump statistic is that if there is no jump in a given time window, then the statistic JS_0 obeys the standard normal distribution. In other words, if the null hypothesis is accepted in month t, that is, there is no jump in the month, then the cumulative positive jump returns ($CJR_{i,t}$) of month t is 0. However, if the test results reject the null hypothesis, it is necessary to further identify the specific date of the jump in month t. The specific identification process is as follows:

Step 1: Let $\{r_1, \cdots, r_{i-1}, r_{median}, r_{i+1}, \cdots, r_N\}$ be daily returns over the interval $[t_1, t_N]$, and record the jump statistic of month t as JS_0.

Step 2: The median r_{median} of the daily return series of the month is used to replace each daily return in the series in turn, and the jump statistics after each replacement are calculated and recorded as JS_i $(i = 1, \cdots, N)$. For instance, when the return on day i is replaced by r_{median}, the new series $\{r_1, \cdots, r_{i-1}, r_{median}, r_{i+1}, \cdots, r_N\}$ will be obtained, and the test statistic is recorded as JS_i.

Step 3: Construct the series $JS_0 - JS_i (i = 1, \cdots, N)$. If $JS_0 - JS_j$ is the highest value of the series, then day j is a jump day.

Step 4: Use r_{median} to replace the daily return on day j, and restart again from Step 1 to identify the new jump day until the jump test statistic accepts the null hypothesis, which means all jumps are identified.

[1] The split share structure reform, implemented in 2005, was completed in 2007, and the reform had an important effect on the financing approach of listed firms (Fang et al., 2018). For this reason, we select 2007 as the starting year. We use high-frequency TAQ (trade and quote) data to calculate the trading behavior of retail investors, and use unique social media data and online search data to calculate investor attention and investor sentiment. Due to the limited availability of intraday data and online social media data, our sample period ends by the end of 2018. The research on market microstructure usually employs subsets of firms and shorter sample periods. For example, Conrad et al. (2015) examine the impact of high-frequency trading on price efficiency using two-year data of the Japanese market. Using no more than 50 component stocks in the German stock market, Dionne and Zhou (2019) study the impact of changes in information environment on price efficiency. Differently, our study uses all eligible stocks in China's stock market to avoid sample selection bias. It also has a relatively long sample period from 01/2007 to 12/2018, including several periods of financial crisis.

[2] https://www.wind.com.cn/en/edb.html

[3] https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/china-stock-market-accounting-research-csmar/

[4] http://www.resset.cn/endatabases

When all jump days in the month t are identified, select the positive returns to accumulate, that is, the cumulative positive jump returns ($CJR_{i,t}$).

3.4. Independent variables

We measure the stock price crash risk by two widely used indicators: the negative conditional skewness (*NCSKEW*) and the down-to-up volatility (*DUVOL*). While stock price crash risk is calculated at the annual frequency in previous studies (e.g., Kim et al., 2011; Yuan et al., 2016; Jang and Kang, 2019; Fu et al., 2021; Chen and Schmidt, 2021), we measure it using the rolling window estimation approach for the next 1 month, 6 months and 12 months. Since market conditions changes through time, measuring the stock price crash risk with data at a relative higher frequency would allow capturing its dynamic variations more accurately. First, we use Eq. (1) to calculate the firm-specific daily return of stock i on day d of month $t + 1$:[5]

$$r_{i,d} = \alpha + \beta_1 r_{m,d-1} + \beta_2 r_{j,d-1} + \beta_3 r_{m,d} + \beta_4 r_{j,d} + \beta_5 r_{m,d+1} + \beta_6 r_{j,d+1} + \varepsilon_{i,d} \tag{1}$$

where $r_{i,d}$ is the return on stock i in day d; $r_{m,d}$ is the value-weighted market return; $r_{j,d}$ is the value-weighted market return of industry j which stock i belongs to. We estimate the firm-specific daily return on stock i in day d of the given time window, that is $w_{i,d} = \ln(1 + \varepsilon_{i,d})$.

Next, based on $w_{i,d}$, we construct the *NCSKEW* and *DUVOL* of stock i in month $t + 1$. The *NCSKEW* is computed as Eqs. (2):

$$NCSKEW_{i,t+1} = -\frac{n(n-1)^{3/2} \sum w_{i,d}^3}{(n-1)(n-2)\left(\sum w_{i,d}^2 \sum_{i,d}^2\right)^{3/2}} \tag{2}$$

where n is the number of trading days of stock i in month $t + 1$.

The $DUVOL_{i,t+1}$ is computed as Eqs.(3):

$$DUVOL_{i,t+1} = ln\left[\frac{(n_u - 1)\sum_{down} w_{i,t}^2}{(n_d - 1)\sum_{up} w_{i,t}^2}\right] \tag{3}$$

where n_u and n_d represent up days and down days, respectively. Larger values of $NCSKEW_{i,t+1}$, and $DUVOL_{i,t+1}$ indicate higher crash risk.[6]

Apart from the impact of positive information shocks in month t on the crashes in month $t + 1$, we also examine the impact of positive information shocks in month t on the crash risk in the next six months ($t + 6$) and next twelve months ($t + 12$). To do so, the firm-specific daily return data from the entire period is used to construct the *NCSKEW* and the *DUVOL* of the next six months ($t + 6$) and next twelve months ($t + 12$).

3.5. Regression model

Using the following model, we identify the impact of positive information shocks on crash risk of stock i:

$$CRASH_{i,t+h} = \beta_0 + \beta_1 CJR_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+h} \tag{4}$$

where t denotes month, and $CRASH_{i,t+h}$ refers to crash risk of stock i in the future h month ($h = 1, 6, 12$), measured by $NCSKEW_{i,t+h}$ and $DUVOL_{i,t+h}$. $CJR_{i,t}$ refers to the positive information shocks of stock i in month t. $CONTROL_{k,i,t}$ is a set of control variables, including one-year lagged $NCSKEW_{i,t}$ and $DUVOL_{i,t}$, mean of firm-specific monthly returns ($Ret_{i,t}$), standard deviation of monthly returns ($Sigma_{i,t}$), financial leverage ($Lev_{i,t}$), book-to-market ratio ($BM_{i,t}$), return on assets ($ROA_{i,t}$), firm size ($Size_{i,t}$), turnover ($Turn_{i,t}$), and absolute accruals ($Accm_{i,t}$). Specific definitions of these variables are presented in the Appendix A. We adopt the month fixed effect and industry fixed effect to estimate Eq. (4). Except for $CJR_{i,t}$, we winsorize all variables at the 1% level in each tail and double cluster in firm and month to correct standard errors.

4. Empirical results

4.1. Descriptive statistics and correlation coefficients

Panel A of Table 1 shows that whatever the time window of crash risk, the mean values of *NCSKEW* and *DUVOL* are less than 0, and they decrease with the extension of the time window. Besides, the standard deviations of these two variables

[5] In order to ensure the sample sufficiency, we use the daily return data of the future three months to get the residuals of Eq. (1). Then, we take the residuals of the future one month (month $t + 1$) to calculate the crash risk indicators of month $t + 1$.

[6] *NCSKEW* contains the characteristics of the second moment (volatility). An increase in *NCSKEW* is also associated with left-skewed return distribution. Therefore, the higher the *NCSKEM*, the greater the stock price crash risk (higher probability of extreme negative return). Chen et al. (2001) propose using *DUVOL* to reduce the interference of the cubic term and a small amount of extreme return rate. The higher the *DUVOL*, the more the return distribution skews to the left, and the higher the stock price crash risk. We also consider the actual stock price decline of 20% as a proxy for the actual price crash. The untabulated results show that positive information shocks have a positive and significant impact on stock price crashes, regardless of price crash thresholds.

Table 1

Descriptive statistics and correlation coefficients. Panel A reports the descriptive statistics of the main variables used in this study. Panel B reports the correlation coefficients, with Pearson correlation coefficients on the above diagonal and Spearman correlation coefficients on the below diagonal. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all variables are winsorized at the 1% level in each tail. And detailed definitions of these variables are reported in Appendix A. ***, **, and * indicates significance at 1%, 5% and 10% levels, respectively.

Panel A: Descriptive statistics

Variables	N	Mean	SD	Q25	Q50	Q75
$NCSKEW_{i,t+1}$	288,688	−0.367	1.121	−1.152	−0.416	0.409
$DUVOL_{i,t+1}$	288,688	−0.344	1.148	−1.121	−0.391	0.391
$NCSKEW_{i,t+6}$	285,811	−0.636	0.403	−1.030	−0.622	−0.233
$DUVOL_{i,t+6}$	285,811	−0.468	0.720	−0.737	−0.483	−0.217
$NCSKEW_{i,t+12}$	266,122	−0.698	0.673	−1.046	−0.688	−0.337
$DUVOL_{i,t+12}$	266,122	−0.509	0.330	−0.729	−0.521	−0.301
CJR	288,688	0.001	0.004	0.000	0.000	0.000
Ret	288,688	−0.001	0.004	−0.002	−0.001	0.002
Sigma	288,688	0.020	0.013	0.013	0.018	0.025
Lev	288,688	0.434	0.222	0.255	0.428	0.599
BM	288,688	0.959	1.101	0.355	0.609	1.097
ROA	288,688	0.027	0.038	0.005	0.019	0.044
Size	288,688	21.871	1.329	20.948	21.683	22.552
Turn	288,688	−0.192	0.645	−0.523	−0.056	0.279
Accm	288,688	−0.001	1.984	−0.025	0.001	0.026

Panel B: Pearson (above diagonal) and Spearman (below diagonal) correlation coefficients

	$NCSKEW_{t+1}$	$DUVOL_{t+1}$	$NCSKEW_{t+6}$	$DUVOL_{t+6}$	$NCSKEW_{t+12}$	$DUVOL_{i,t+12}$	*CJR*	*Ret*	*Sigma*	*Lev*	*BM*	*ROA*	*Size*	*Turn*	*Accm*
$NCSKEW_{t+1}$		0.939***	0.174***	0.201***	0.116***	0.151***	0.050***	0.069***	0.118***	0.001	−0.016***	0.007***	−0.007***	0.093***	0.001
$DUVOL_{t+1}$	0.947***		0.144***	0.187***	0.097***	0.141***-	0.051***	0.072***	0.116***	0.002	−0.012***	0.004**	0.005**	0.097***	−0.001
$NCSKEW_{t+6}$	0.175***	0.151***		0.861***	0.650***	0.610***	0.015***	0.013***	0.063***	−0.019***	−0.098***	0.017***	−0.060***	0.023***	0.001
$DUVOL_{t+6}$	0.194***	0.183***	0.877***		0.551***	0.681***	0.004**	0.017***	0.027***	−0.005**	−0.085***	0.013***	−0.067***	0.022***	−0.002
$NCSKEW_{t+12}$	0.125***	0.108***	0.680***	0.601***		0.839***	0.015***	0.005**	0.076***	−0.037***	−0.116***	0.026***	−0.058***	0.006***	0.005**
$DUVOL_{t+12}$	0.148***	0.140***	0.631***	0.686***	0.874***		0.003	−0.012***	−0.026***	−0.012***	−0.114***	0.018***	−0.067***	0.000	−0.002
CJR	0.046***	0.046***	0.025***	0.021***	0.029***	0.023***		0.089***	0.085***	−0.002	−0.031***	0.001	−0.024***	0.090***	0.001
Ret	0.101***	0.104***	0.026***	0.027***	−0.131***	0.050***	0.151***		−0.353***	0.004**	0.008***	−0.003*	0.004**	0.217***	0.011***
Sigma	0.152***	0.144***	0.104***	0.048***	0.136***	−0.045***	0.054***	0.080***		0.003	−0.134***	−0.002	−0.154***	0.155***	−0.019***
Lev	−0.020***	−0.018***	−0.058***	−0.078***	−0.075***	−0.103***	−0.014***	0.001	−0.024***		0.021***	−0.431***	−0.027***	0.004**	−0.166***
BM	−0.043***	−0.035***	−0.168***	−0.148***	−0.214***	−0.196***	−0.036***	0.013***	−0.270***	0.576***		−0.026***	0.575***	0.019***	0.001
ROA	0.033***	0.030***	0.098***	0.101***	0.115***	0.124***	−0.011***	−0.022***	−0.024***	−0.336***	−0.277***		0.053***	−0.005***	0.030***
Size	−0.011***	−0.008***	−0.077***	−0.082***	−0.087***	−0.089***	−0.016***	−0.003	−0.224***	0.417***	0.615***	0.008***		0.025***	0.004*
Turn	0.106***	0.107***	0.033***	0.033***	0.012***	0.006***	0.101***	0.295***	0.203***	0.015***	0.035***	−0.011***	0.035***		0.003
Accm	−0.002	−0.002	−0.018***	−0.013***	−0.012***	−0.012***	−0.003	0.003*	−0.007***	−0.003*	0.035***	−0.011***	0.007***	−0.005***	

are large, especially in the case of $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ which are 1.121 and 1.148, respectively, suggesting that there is a big crash risk difference between individual stocks. It also indicates that the cumulative positive jump returns ($CJR_{i,t}$) are still 0 in the 75th percentile, but the months with jumps only account for 15.75% of the total sample,[7] which fully reflects the rarity of stock price jumps. Hence, positive information shock does not exist for all stocks, and it has the feature of a right fatted-tail distribution. Due to the relative scarcity of positive information shocks, retail investors will have relative behavior deviation in the face of the positive information shocks. The statistical results of control variables are all in a reasonable range.

Panel B of Table 1 shows that regardless of the length of the time window, the correlations of the two crash risk indicators under the same data frequency are greater than 0.8, and both are significant at the level of 1%, which means these two indicators are highly positively correlated and have good consistency in the measurement of crash risk. The longer the time window, the lower the correlation between the two crash risk proxies. The cumulative positive jump returns ($CJR_{i,t}$) has significant positive correlation with the crash risk of any time window in the future, and the correlation degree in the short term is more significant than that in the long term. The correlation coefficients between the control variables are in a reasonable range, so the multicollinearity problem does not need to be considered. Meanwhile, the correlation coefficients between the control variables and the crash risk indices are significant, which indicates that these control variables may affect the crash risk. Therefore, it is necessary to control these potential factors in the regression models to obtain reliable results.

4.2. Basic results

The main results are shown in Table 2. The time window is the next month, that is $h = 1$ in Eq. (4). The coefficients on $CJR_{i,t}$ are positive at the 1% level, no matter how crash risk is measured and whether control variables are added. Thus, the higher the positive information shocks, the higher the crash risk in the next month, supporting *Hypothesis 1b* that positive information shocks exacerbate stock price crash risk.

Moreover, the coefficients on $NCSKEW_{i,t}$ are positive, indicating that current crash risk aggravates the crash risk in the next month, which is in line with Cheng et al. (2020). The coefficients on $Ret_{i,t}$ and $Sigma_{i,t}$ are significantly positive, showing that the higher the firm-specific returns, the more intense the volatility, and the greater the crash risk in the future, which is consistent with Xu et al. (2014). Besides, the coefficient on $LEV_{i,t}$ is also significantly positive, meaning that the higher the debt level, the higher the crash risk in the next month, confirming the conclusion of Wen et al. (2019). The coefficients on $BM_{i,t}$ are negative, as in Harvey and Siddique (2000). The coefficients on $Accm_{i,t-1}$ are positive, which suggests that stocks with higher degrees of accrual basis manipulation are likely to crash. Kim et al. (2011b) reaches the same conclusion.

We further extend the time window to the next six and twelve months, that is, $h = 6$ and $h = 12$ in Eq. (4). The regression results are shown in Table 3. Panel A of Table 3 indicates that when the time window includes the next six months, the coefficients on $CJR_{i,t}$ are positive at the 1% level. Although the significance of coefficients decreases after control variables are added, they remain significant at the 5% level. Even if other known characteristics affecting the crash risk are controlled, the significant impact of positive information shocks on the crash risk is still there. When the time window is extended to the next 12 months, the results shown in Panel B are similar to those in Panel A and point out that positive information shocks are positively associated with crash risk in the next 12 months.

According to Tables 2 and 3, positive information shocks aggravate the crash risk and the impact in the short term is stronger than that in the long term. Specifically, the influence of positive information shocks on crash risk in the next month is significantly stronger than that in the next 6 months and in the next 12 months, and the impact in the next 6 months is also significantly stronger than that in the next 12 months. In the following discussion, we will focus on the crash risk of the month (t + 1) to capture the dynamic change of crash risk.

4.3. Robustness test

4.3.1. Additional control variables

Following Cheng et al. (2021), we added liquidity (*Liquidity*), volatility (*Volatility*), institutional shareholding (*INST*), and managerial shareholding (*Mshare*) in the baseline model to avoid the potential endogeneity problem caused by omitted variables. We then re-estimate the new regression model. The results in Panel A of Table 4 show that when *NCSKEW* is used to measure the stock price crash risk, the coefficient of *CJR* is still significantly positive at the 1% level after controlling the variables such as liquidity, volatility, the shareholding ratio of institutional investors, and the shareholding ratio of management successively. We obtain the same result when these control variables are simultaneously introduced in the regression. Panel B reports similar findings when stock price crash risk is proxied by *DUVOL*.

4.3.2. Alternative measures of positive information shock

We now examine the result sensitivity to other types of positive information shock measurement. First, a jump that will be observed in a day (i) during a month (j) might be diluted by the trend of all the month. In order to deal with the above

[7] The 85% quantile of CJR is 0.01; The 90% quantile is 0.05; The 95% quantile is 0.09; The 99% quantile is 0.19.

Table 2

The impact of CJR on crash risk over the next month. This table presents the impact of CJR on crash risk over the next month while controlling for industry and month fixed effects. The regression model is as follows:

$$CRASH_{i,t+1} = \beta_0 + \beta_1 CJR_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+1}$$

where $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ are the measure of $CRASH_{i,t+1}$ of firm *i* in month *t + 1*, $CJR_{i,t}$ denotes the CJR measure, and $CONTROL_{k,i,t}$ is a set of control variables. Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column (1) and Column (3) employ the fixed effects regressions without controls, while Column (2) and Column (4) employ the fixed effects regressions with controls.

Dependent variable =	$NCSKEW_{i,t+1}$		$DUVOL_{i,t+1}$	
	(1)	(2)	(3)	(4)
CJR	7.194***	3.456***	7.515***	3.621***
	(6.761)	(5.715)	(7.237)	(6.299)
NCSKEW		0.020**		
		(2.268)		
DUVOL				0.014
				(1.406)
Ret		2.444***		2.484***
		(9.559)		(9.298)
Sigma		1.712***		1.746***
		(9.694)		(9.883)
Lev		0.002***		0.002***
		(2.734)		(3.021)
BM		−0.013		−0.009
		(−1.500)		(−1.088)
ROA		0.027		0.016
		(1.282)		(0.877)
Size		0.022**		0.021**
		(2.491)		(2.209)
Turn		0.001***		0.001***
		(4.387)		(4.544)
Accm		0.002***		0.001
		(3.557)		(1.465)
Constant	−0.631***	−1.565***	−0.708***	−1.622***
	(−21.905)	(−7.688)	(−25.294)	(−7.611)
Month fixed effects	YES	YES	YES	YES
Industry fixed effects	YES	YES	YES	YES
Adjusted R^2	0.049	0.082	0.048	0.082

potential bias, the new jump test is performed using daily return observations over a three-month rolling window. Once the above jump test rejects the null hypothesis of no jumps in a given three-month window, we proceed to identify those days with stock price jumps following a sequential procedure. Let $\{r_1, \cdots, r_{i-1}, r_{i+1}, \cdots, r_N\}$ be daily returns over the interval $[t_1, t_N]$. The null hypothesis of the jump statistic is that if there is no jump in a given three month rolling window, then the statistic JS_0 obeys the standard normal distribution. If the jump test does not reject the null hypothesis of no jumps, we move to the next three-month window. In other words, if the null hypothesis is not rejected in month (t-2, t), that is, there is also no jump in the month t, then the cumulative positive jump returns ($CJR_{i,t}$) of month t is 0. However, if the test results reject the null hypothesis, it is necessary to further identify the specific date of the jump in the three-month window. According to the same method introduced in this paper, we can detect the positive jump in month t, and then sum up the relevant positive jumps returns as the new measurement index *CJR1*.

In addition, we also adopt a new jump test methodology to avoid the dependence of the results on the jump test method. Following Lee and Mykland (2008) and Jawadi et al. (2019), we use intraday jump tests according to the corresponding 5-minute returns. This test examines the null hypothesis of continuity across the price dynamics against the alternative hypothesis of intraday jumps. We accumulate all positive intraday jump returns in the month *t* as a new measure *CJR2*.

Finally, following Yu et al. (2020), we also use the cumulative positive large jump volatility as a new measurement index *CJR3* Anderson et al. (2007). propose that realized volatility can be decomposed into continuous volatility and jump volatility Yu et al. (2020). further decompose the jump volatility into large and small jump volatility. They point out that large jump volatility is mainly caused by information shocks, while small jump volatility is caused by short-term liquidity changes or strategic trading of stocks. We follow Yu et al. (2020) and decompose jump volatility into positive and negative ones for each trading day. Then, using the threshold division method, the positive jump volatility is further decomposed into large and small positive jump volatility. This decomposition process relies on the corresponding 5-minute returns. We accumulate all large positive jump volatility in the month *t* as a new measurement *CJR3*.

Table 3
The impact of CJR on crash risk over the next six and twelve months. This table presents the impact of CJR on crash risk over the next six and twelve months while controlling for industry and month fixed effects. The regression model is as follows:

$$CRASH_{i,t+h} = \beta_0 + \beta_1 CJR_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+h}$$

where $NCSKEW_{i,t+h}$ and $DUVOL_{i,t+h}$ are the measure of $CRASH_{i,t+h}$ of firm i in month $t + h$, $h = 6, 12$, $CJR_{i,t}$ denotes the CJR measure, and $CONTROL_{k,i,t}$ is a set of control variables. Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). t-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Panel A reports the effect of CJR on crash risk over the next six months ($h = 6$), while Panel B reports the effect of CJR on crash risk over the next twelve months ($h = 12$). Column (1) and Column (3) employ the fixed effects regressions without controls, while Column (2) and Column (4) employ the fixed effects regressions with controls.

Panel A: The impact of CJR on crash risk over the next six months				
Dependent variable =	$NCSKEW_{i,\,t+6}$		$DUVOL_{\,i,\,t+6}$	
	(1)	(2)	(3)	(4)
CJR	1.565***	0.691**	0.938***	0.489**
	(3.657)	(2.060)	(3.803)	(2.353)
Constant	−0.593***	−0.941***	−0.560***	−0.607***
	(−18.751)	(−6.457)	(−29.993)	(−7.121)
Controls	NO	YES	NO	YES
Month fixed effects	YES	YES	YES	YES
Industry fixed effects	YES	YES	YES	YES
Adjusted R^2	0.064	0.081	0.115	0.126

Panel B: The impact of CJR on crash risk over the next twelve months				
Dependent variable =	$NCSKEW_{i,\,t+12}$		$DUVOL_{\,i,\,t+12}$	
	(1)	(2)	(3)	(4)
CJR	1.089***	0.507*	0.492***	0.212*
	(3.501)	(1.908)	(3.812)	(1.877)
Constant	−0.585***	−0.833***	−0.545***	−0.565***
	(−14.896)	(−5.158)	(−26.543)	(−6.798)
Controls	NO	YES	NO	YES
Month fixed effects	YES	YES	YES	YES
Industry fixed effects	YES	YES	YES	YES
Adjusted R^2	0.059	0.075	0.132	0.146

As shown in Table 5, our baseline results are not affected by changing the time window of the jump test or changing the jump detection methodology.

4.4. Endogeneity issues

4.4.1. Multiple fixed effect model

Our results pointed to a positive relationship between positive information shocks and crash risk. To reduce the potential impact of relevant missing variables, we follow Kim et al. (2011b) and introduce the firm fixed effect and time fixed effect in Eq. (4), which captures unpredictable firm-specific factors affecting the crash risk. The results reported in Panel A of Table 6 show that our baseline results remain unchanged when we use the firm fixed-effect model since the coefficients on $CJR_{i,t}$ of $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ are both positive and significant at the 1% level.

To further solve the endogeneity issues, we also control the time-varying industry fixed effect. The results in Panel B of Table 6 show that the coefficients associated with $CJR_{i,t}$ of $NCSKEW_{i,t+1}$ *and* $DUVOL_{\,i,t+1}$ are all significantly positive at the 1% level, which verifies the robustness of our baseline results.

4.4.2. Instrumental variable method

We follow Kubick and Lockhart (2021) and use the average *CJR* with the same industry in the month t as IV to rule out the impact of omitted variables. The two-stage least squares (2SLS) regressions are carried out to examine the effect of *CJR* on the crash risk. The first-stage regression results, reported in Column (1) in Table 7, shows that the *IV* is positive and statistically significant. The results in Column (2) and Column (3) are in line with our baseline results, indicating that the 2SLS results support our findings.

Table 4

Additional control variables. This table reports the results controlling for possibly omitted variables. In Panel A, the dependent variable is $NCSKEW_{i,t+1}$ while in Panel B, the dependent variable is $DUVOL_{\,i,t+1.}$ In addition to the control variables in basic model, we add *Liquidity* to proxy for stock liquidity in Column (1), *Volatility* to proxy for stock volatility in Column (2), *INST* to proxy for institutional shareholding in Column (3), *Mshare* to proxy for managerial shareholding in Column (4), and all the above control variables are included in Column (5). Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A					
Dependent variable =	$NCSKEW_{i,t+1}$				
	(1)	(2)	(3)	(4)	(5)
CJR	3.166***	2.750***	3.273***	3.187***	2.679***
	(6.000)	(5.227)	(6.063)	(5.951)	(5.060)
Liquidity	−0.212*				−0.650***
	(−1.795)				(−5.390)
Volatility		2.193***			2.707***
		(8.720)			(11.265)
INST			−0.097***		−0.139***
			(−3.432)		(−5.027)
Mshare				0.017	0.074*
				(0.421)	(1.814)
Constant	−1.761***	−2.252***	−1.867***	−2.004***	−1.956***
	(−8.961)	(−11.287)	(−9.267)	(−9.949)	(−9.812)
Controls	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Adjusted R^2	0.092	0.096	0.094	0.095	0.102
Panel B					
Dependent variable =	$DUVOL_{\,i,t+1}$				
	(1)	(2)	(3)	(4)	(5)
CJR	3.336***	2.884***	3.446***	3.885***	2.830***
	(6.574)	(5.783)	(6.718)	(7.214)	(5.664)
Liquidity	−0.161*				−0.604***
	(−1.698)				(−4.764)
Volatility		2.363***			2.851***
		(8.626)			(10.827)
INST			−0.099***		−0.134***
			(−3.324)		(−4.567)
Mshare				0.004	0.056
				(0.105)	(1.384)
Constant	−1.752***	−2.325***	−1.881***	−2.012***	−2.031***
	(−8.380)	(−10.999)	(−8.924)	(−9.580)	(−9.705)
Controls	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Adjusted R^2	0.090	0.095	0.094	0.093	0.100

4.4.3. GMM-dynamic panel data model

The stock price crash risk of listed firms is vulnerable to the past crash risk, and the dynamic panel model can control the impact of lag factors on current factors. Following Cui et al. (2021), we use the dynamic panel model with multi-periods' lag stock price crash risk to re-estimate our model to control for the impact of relevant lag factors. The GMM estimation results of Table 8 confirm those of our baseline models.

4.4.4. Propensity score matching method (PSM)

Since there may be other potential unobservable factors which cause positive information shock, we use PSM method for the regression to further eliminate the impact of other unobservable factors on the results. Following Hu et al. (2020), we set the firms with non-zero *CJR* in the month *t* as the treatment group and set the samples with zero CJR as the control group. We then estimate a logistic model to predict the probability of being the treated firms on a series of control variables (i.e., *INST, Mshare, Size, ROA, Lev, BM, Ret*) used in the above analysis. Column (1) of Table 9 shows the matching results of the first step. Column (2) and Column (3) report the regression results with matching sample. We find the coefficient of CJR is still positive at 1% level, no matter using $NCSKEW_{i,t+1}$ or using $DUVOL_{\,i,t+1}$ to proxy crash risk.

Table 5

Alternative measures of positive information shocks. The table reports the estimation results of alternative measurement of positive information shocks. As for CJR1, the new jump test is performed using daily return observations over a three-month rolling window. Using the method introduced in Section 3.3, we can detect the positive jump in month *t*, and then sum up the relevant positive jumps returns as the new measurement index *CJR1*. Following Lee and Mykland (2008) and Jawadi et al. (2019), we accumulate all positive intraday jump returns in the month t as a new measure *CJR2*. Following Yu et al. (2020), we accumulate all large positive jump volatility in the month t as a new measurement *CJR3*.

Dependent variable =	$NCSKEW_{i,t+1}$			$DUVOL_{i,t+1}$		
	(1)	(2)	(3)	(4)	(5)	
CJR1	0.133***			0.141***		
	(4.535)			(4.383)		
CJR2		0.223***			0.240***	
		(5.752)			(5.647)	
CJR3			0.089***			0.041***
			(2.714)			(3.236)
NCSKEW	0.020***	0.004*	0.015**			
	(3.073)	(1.678)	(2.229)			
DUVOL				0.012*	0.003	0.007
				(1.754)	(0.458)	(0.994)
Ret	1.435***	1.853***	2.073***	1.398***	1.842***	2.073***
	(5.575)	(7.912)	(8.559)	(4.992)	(7.213)	(7.793)
Sigma	1.774***	1.423***	2.388***	1.728***	1.343***	2.384***
	(3.428)	(9.816)	(9.365)	(3.317)	(8.809)	(8.344)
Lev	−0.037	−0.039	−0.024	−0.047*	−0.050*	−0.034
	(−1.363)	(−1.406)	(−0.848)	(−1.745)	(−1.795)	(−1.221)
BM	−0.001	−0.003	−0.002	0.003	0.002	0.002
	(−0.133)	(−0.232)	(−0.212)	(0.295)	(0.208)	(0.145)
ROA	1.089***	1.145***	0.901***	0.959***	1.018***	0.775***
	(5.828)	(6.166)	(4.856)	(4.868)	(5.184)	(3.934)
Size	0.030***	0.039***	0.026***	0.029***	0.038***	0.026***
	(3.459)	(4.384)	(3.022)	(3.173)	(4.111)	(2.842)
Turn	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(3.438)	(3.482)	(4.824)	(3.594)	(4.043)	(4.709)
Accm	−0.011	−0.018	−0.012	−0.014	−0.022	−0.015
	(−0.230)	(−0.379)	(−0.261)	(−0.292)	(−0.447)	(−0.299)
Constant	−1.896***	−2.265***	−1.801***	−1.942***	−2.339***	−1.875***
	(−9.704)	(−10.365)	(−9.178)	(−9.447)	(−10.079)	(−9.065)
Month FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Adjusted R^2	0.098	0.096	0.091	0.097	0.095	0.090

5. An economic linkage from positive information shocks to crash risk

5.1. Transmission mechanism: the role of retail investors

Our baseline results show that positive information shocks have a significantly positive impact on the crash risk. This section further studies the potential transmission path of the relationship.

In *Hypothesis 1a*, we expect that investors who are attracted by positive information shocks may pay more attention to those stocks with positive shocks and make efforts to search for related information about them, playing the monitoring role in mitigating information asymmetry and leading to lower crash risk. However, regression results suggest that positive information shocks cannot reduce the crash risk, but on the contrary, they aggravate the crash risk, verifying *Hypothesis 1b*. In *Hypothesis 1b*, we assume that retail investors play a necessary role in the positive relationship.

We argue that the impact of positive information shocks will attract the attention of retail investors (Vozlyublennaia, 2014; Yao et al., 2021). The latter is closely related to investors' sentiment changes and trading behavior (Shleifer and Summers, 1990; Campbell and Kyle, 1993; Lakonishok et al., 1994; Barberis et al., 1998; Jawadi et al., 2018). Studies such as Loewenstein (2001), Böhm and Chiarella (2005), and Ftiti et al. (2016) find that retail traders are likely to affect stock prices and investors' decision-making process based on unpredictable changes in sentiment. When retail investors are attracted by positive information shocks, abnormal retail attention, over-optimistic sentiment, and excessive trades may play a role simultaneously, leading to overvalued stock prices. Moreover, aggressive retail investor behavior will cause strong limits of arbitrage, hinder timely price correction, and further trigger corresponding speculation (Yao et al., 2019). This retail investor behavior has no support from the firms' fundamentals, and relevant mispricing will exacerbate the stock crash risk in the future.

Table 6
Multiple-fixed effect model. This table presents the impact of CJR on crash risk over the next month while controlling for firm and month fixed effects (in Panel A) and controlling for firm and time-varying industry fixed effects (in Panel B). Column (1) and Column (3) employ the fixed effects regressions without controls, while Column (2) and Column (4) employ the fixed effects regressions with controls. Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Firm fixed effect model				
Dependent variable =	$NCSKEW_{i,t+1}$		$DUVOL_{i,t+1}$	
	(1)	(2)	(3)	(4)
CJR	7.097***	3.435***	7.442***	3.616***
	(6.449)	(5.707)	(6.984)	(6.321)
Constant	−0.459***	−0.961***	−0.549***	−1.016***
	(−16.757)	(−3.623)	(−19.163)	(−3.673)
Controls	NO	YES	NO	YES
Month fixed effects	YES	YES	YES	YES
Firm fixed effects	YES	YES	YES	YES
Adjusted R^2	0.061	0.082	0.058	0.082
Panel B: Time-varying industry fixed effect model				
Dependent variable =	$NCSKEW_{i,t+1}$		$DUVOL_{i,t+1}$	
	(1)	(2)	(3)	(4)
CJR	6.716***	3.182***	7.035***	3.377***
	(7.040)	(3.105)	(7.585)	(4.741)
Constant	−0.606***	−1.706***	−0.617***	−1.765***
	(−7.754)	(−8.272)	(−7.835)	(−7.755)
Controls	NO	YES	NO	YES
Industry FE*Year FE	YES	YES	YES	YES
Firm fixed effects	YES	YES	YES	YES
Adjusted R^2	0.012	0.052	0.012	0.052

From the perspective of retail investor behavior,[8] we further explore the potential path of positive information shocks affecting crash risk according to three aspects: (i) retail investors may pay abnormal attention to the positive jumps of stock price, which promotes the rise of stock price, causes it to deviate from the fundamental value, and finally aggravates the future crash risk; (ii) retail investors may be over-optimistic about the positive information shocks, and reflect the abnormal sentiment onto the stock price, leading to the overvaluation of the stock price and eventually aggravating the crash risk; and (ii) the positive jumps of stock prices may stimulate the net buying behavior of retail investors, and excessive behavior bias eventually aggravates the crash risk Fig. 1. illustrates the above-mentioned mechanism path.

A two-step regression method is used to verify the above three transmission paths. Specifically, the first step is to investigate the relationship between positive information shocks and mediating variables, and the second step is to study the impact of mediating variables on future crash risk. Therefore, we construct Eqs. (5) and (6) in turn:

$$Channel_{i,t} = \beta_0^1 + \beta_1^1 CJR_{i,t} + \sum_k \gamma_k^1 CONTROL_{k,i,t}^1 + \varepsilon_{i,t} \tag{5}$$

$$CRASH_{i,t+1} = \beta_0^2 + \beta_1^2 Middle_{i,t} + \sum_k \gamma_k^2 CONTROL_{k,i,t}^2 + \varepsilon_{i,t+1} \tag{6}$$

where $Channel_{i,t}$ refers to the mediating variables, e.g., retail investor abnormal attention $Abn\_SVI_{i,t}$, abnormal sentiment $Abn\_Post_{i,t}$, and retail net buy behavior $SOIB_{i,d}^{retail}$. It is worth noting that all variables in Eq. (5) are in the same period, while the mediating variables and control variables in Eq. (6) lag for one period.

5.1.1. Retail investor abnormal attention

Following Zhang et al. (2014), we construct a retail investor abnormal attention index ($Abn\_SVI_{i,t}$) of stock i in month t. We adopt Baidu search volume to proxy for retail attention ($SVI_{i,d}$) of stock i on day d. $Abn\_SVI_{i,d}$ can be obtained by Eq. (7), where $average(SVI_{i,(d-90,d-31)})$ is the average value of the SVI of stock i from day d-90 to day d-31. Finally, $Abn\_SVI_{i,t}$ is measured as the sum of $Abn\_SVI_{i,d}$ on the following day of the positive jump day in month t.

Table 7

Instrumental variable method. This table reports the results from the 2SLS regression using instrumental variable. This table reports the results from the 2SLS regression using instrumental variable. The variable *IV* is calculated as the average value of CJR of firms in the same industry in the same month. In the first stage, the dependent variable is $CJR_{i,t.}$ In the second stage, the dependent variables are $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1.}$ Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable	First_Stage $CJR_{i,t}$ (1)	Second_Stage $NCSKEW_{i,t+1}$ (2)	Second_Stage $DUVOL_{i,t+1}$ (3)
$\overline{\overline{P}}red.$ *CJR*		3.174***	3.209***
		(3.330)	(3.206)
IV	0.931***		
	(9.960)		
NCSKEW	0.001**	0.014***	
	(2.506)	(3.736)	
DUVOL			0.006**
			(2.167)
Ret	0.022***	1.732***	1.721***
	(3.322)	(8.874)	(8.020)
Sigma	0.009***	2.242***	2.234***
	(9.802)	(6.709)	(6.635)
Lev	−0.001***	−0.019	−0.028
	(−3.386)	(−0.423)	(−0.157)
BM	−0.001	−0.003	0.001
	(−0.885)	(−1.006)	(0.397)
ROA	0.004***	0.854***	0.712***
	(3.946)	(3.509)	(3.295)
Size	0.001	0.027***	0.026***
	(0.289)	(2.684)	(3.416)
Turn	0.001***	0.001***	0.001***
	(6.356)	(3.078)	(3.430)
Accm	−0.001	−0.011	−0.015
	(−0.864)	(−0.339)	(−0.445)
Constant	−0.003***	−1.809***	−1.849***
	(−7.223)	(−8.374)	(−8.790)
Month FE	YES	YES	YES
Industry FE	YES	YES	YES
Adjusted R^2	0.072	0.083	0.082

Table 8

Dynamic panel data model-GMM. This table reports the dynamic panel model estimated using the GMM method. Referring to Cui et al. (2021), we use the dynamic panel model to control the multi-period lag stock price crash risk. AR (1) and AR (2) are used to verify whether there is first-order and second-order sequence correlation in GMM estimation, and Hansen test and Difference-in-Hansen tests are used to verify whether the instrumental variables in horizontal and differential forms in GMM estimation are reasonable. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable =	$NCSKEW_{i,t+1}$ (2)	$DUVOL_{i,t+1}$ (4)
CJR	3.609***	3.579***
	(3.616)	(3.804)
NCSKEW (−1)	0.011***	
	(3.072)	
NCSKEW (−2)	0.015	
	(0.299)	
DUVOL (−1)		0.009***
		(3.142)
DUVOL (−2)		0.002
		(0.460)
Controls	YES	YES
Month FE	YES	YES
Industry FE	YES	YES
AR (1) Test (P value)	0.000	0.000
AR (2) Test (P value)	0.228	0.320
Hansen Test	0.448	0.425
Difference-in-Hansen Tests	0.510	0.461

Table 9

Propensity score matching model (PSM). This table reports the regression results after PSM. We set the treatment group and the control group according to whether the firms have non-zero CJR in the month t. If the firm has non-zero CJR in the month t, we set it as the treatment group (*CJR Dummy=1*). We match the treatment group one-to-one according to a series of variables, including *INST, Mshare, Size, ROA, LEV, BM* and *Ret*. Column (1) reports the matching results. Column (2) and (3) report the regression results with matching sample, using $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ to proxy for crash risk. Detailed definitions of all variables are shown in Appendix A. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable =	*CJR Dummy* (1)	$NCSKEW_{i,t+1}$ (2)	$DUVOL_{i,t+1}$ (3)
CJR		3.189***	3.354***
		(6.042)	(6.625)
INST	0.259***		
	(9.661)		
Mshare	0.176***		
	(5.630)		
Size	−0.008	0.027***	0.026***
	(−1.257)	(3.151)	(2.845)
ROA	−1.327***	0.899***	0.757***
	(−8.485)	(4.860)	(3.885)
Lev	0.084***	−0.023	−0.032
	(2.597)	(−0.854)	(−1.191)
BM	−0.118***	−0.004	0.001
	(−3.780)	(−0.335)	(0.087)
Ret	8.787***	2.017***	2.014***
	(6.013)	(8.445)	(7.698)
NCSKEW		0.015**	
		(2.246)	
DUVOL			0.007
			(1.008)
Sigma		2.360***	2.351***
		(9.354)	(8.285)
Turn		0.001***	0.001***
		(4.676)	(4.760)
Accm		−0.013	−0.016
		(−0.270)	(−0.328)
Constant	−1.523***	−1.832***	−1.872***
	(−12.292)	(−9.490)	(−9.195)
Month FE	YES	YES	YES
Industry FE	YES	YES	YES
Pseudo R2/Adjusted R^2	0.022	0.092	0.091

Panel A of Table 10 shows the results of $Abn\_SVI_{i,t}$ as the mediating variable.

$$Abn\_SVI_{i,d} = \frac{SVI_{i,d} - average\left(SVI_{i,(d-90,d-31)}\right)}{average\left(SVI_{i,(d-90,d-31)}\right)} \tag{7}$$

The coefficient on $CJR_{i,t}$ of $Abn\_SVI_{i,t}$ is positive at the 1% level in Column (1), suggesting that the positive information shocks have aroused the abnormal retail investor attention. The positivity of the coefficients on $Abn\_SVI_{i,t}$ of $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ in Column (2) and Column (3) indicates that the retail abnormal investor attention aggravates the crash risk.

The evidence shows that positive information shocks arouse the abnormal retail investor attention, but what comes with the abnormal attention is not the weakening of information asymmetry and the reduction of crash risk, but the further increase of crash risk. Therefore, we can draw the conclusion that abnormal retail investor attention is the intermediate transmission path from the positive information shocks to crash risk.

5.1.2. Abnormal retail investor sentiment

Huang et al. (2016) document that the stock community Guba Eastmoney Forum provides a network communication platform for investors, especially for retail investors to conduct real-time market comments and stock information exchange. We use the number of abnormal posts on Guba ($Abn\_Post_{i,t}$) to measure the abnormal retail investor sentiment. First, we use the number of positive posts on Guba to measure the optimism of retail investors ($gb_{i,d}$) of stock *i* on day *d*. Furthermore,

[8] We also employ the large trades and analyst upgrades within the month after the positive information shocks to measure the change in institutional sentiment and trading behavior. According to the untabulated results, there is no significant relationship between crash risk and the variation in institutional sentiment and trading.

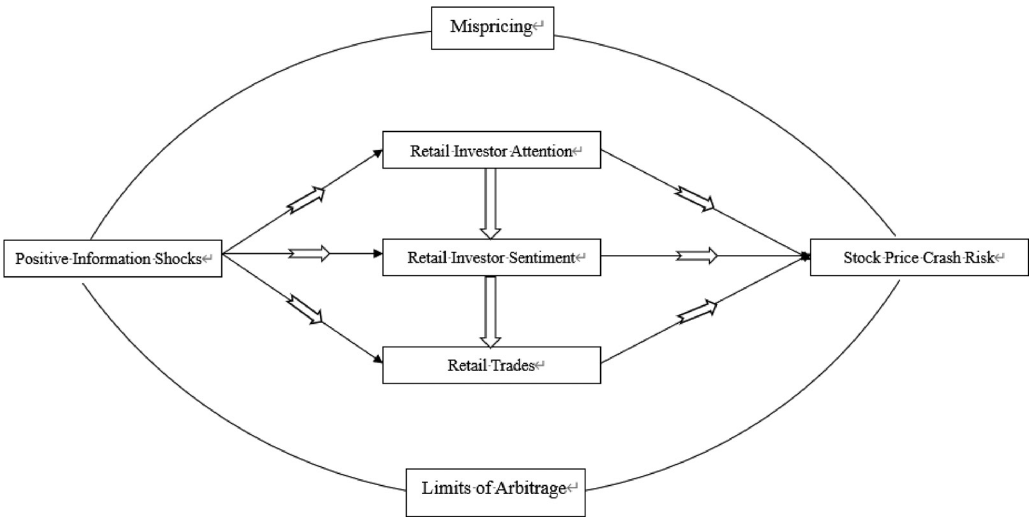

Fig. 1. Structure figure of channel path. The potential path of positive information shocks affecting crash risk from three aspects: (1) retail investors may pay abnormal attention to the positive jumps of stock price, which promotes the rise of stock price, causes it to deviate from the fundamental value, and finally aggravates the future crash risk, (2) retail investors may be over optimistic about the positive information shocks, and reflect the abnormal sentiment onto the stock price, leading to the overvaluation of the stock price and eventually aggravating the crash risk, and (3) the positive jumps of stock prices may stimulate the net buying behavior of retail investors, and excessive behavior bias eventually aggravates the crash risk. We provide the structure figure of channel path as follows.

we can get the abnormal positive posts $Abn\_Post_{i,d}$ by Eq. (8), where $average(gb_{i,(d-90,d-31)})$ is the average value of *gb* of stock *i* from day *d-90* to day *d-31*. Finally, $Abn\_Post_{i,t}$ is measured as the sum of $Abn\_Post_{i,d}$ on the following day of the positive jump day in month *t*.

$$Abn\_Post_{i,d} = \frac{gb_{i,d} - average\left(gb_{i,(d-90,d-31)}\right)}{average\left(gb_{i,(d-90,d-31)}\right)} \tag{8}$$

Panel B of Table 10 shows the results of $Abn\_Post_{i,t}$ as the mediating variable. The coefficient on $CJR_{i,t}$ is positive in Column (1) of Panel B, which shows that the positive information shocks induce over-optimistic retail investor sentiment. Column (2) and Column (3) report the results of the second step. The coefficients on $Abn\_Post_{i,t}$ are positive, representing that abnormal retail sentiment also has a positive effect on the crash risk. Therefore, we can draw the conclusion that abnormal retail investor sentiment is the transmission path of the exacerbating effect of positive information shocks on crash risk.

5.1.3. Retail investor net buys

We now turn to test whether retail investors affected by positive information shocks carry out abnormal trading activities, ultimately increasing the stock price crash risk Lee and Radhakrishna (2000). propose that small-scale transactions are mainly initiated by retail investors. Thus, in our study, the transaction orders of less than 100,000 yuan are used to proxy for the retail investor trading behavior. Following Ng and Wu (2007), we construct the retail investor net buying behavior $SOIB_{i,t}$. First of all, we construct the retail investor initiative net buying index $SOIB^{retail}_{i,d}$, as shown in Eq. (9), where $buy^{retail}_{i,d}$ and $sell^{retail}_{i,d}$ correspond to the purchase transaction volume and sale transaction volume of retail investors to stock *i* on day *d*, respectively. The monthly net buys ($SOIB_{i,t}$) are measured as the sum of $SOIB_{i,d}$ on the following trading day of the positive jump day in month *t*. The results of $SOIB_{i,t}$ as the mediating variable are shown in Panel C of Table 10.

$$SOIB^{retail}_{i,d} = \frac{buy^{retail}_{i,d} - sell^{retail}_{i,d}}{buy^{retail}_{i,d} + sell^{retail}_{i,d}} \tag{9}$$

The coefficient on $CJR_{i,t}$ is positive at the 5% level in Column (1) of Panel C, indicating that retail investors take the initiative to buy the stocks with positive information shocks. The coefficients on $SOIB_{i,t}$ are significantly positive in Column (2) and Column (3), representing that the active net buying aggravates the crash risk. The above results show that "retail net buys" are also important channels for the exacerbating effect of positive information shocks on crash risk.

In summary, positive information shocks attract abnormal retail investor attention and intensify their optimistic sentiment and net buying behavior which promote the crash risk. Therefore, abnormal retail investor attention, abnormal sentiment and net buys are the intermediate transmission paths of the positive information shocks aggravating the crash risk.

Table 10

Channel test: retail investor attention, retail investor sentiment, and retail net-buys behavior. This table reports the regression results of retail investor attention, retail investor sentiment, and retail net-buys behavior as the channel mechanisms while controlling for industry and month fixed effects. The regression model is as follows:

$$Middle_{i,t} = \beta_0^1 + \beta_1^1 CJR_{i,t} + \sum_k \gamma_k^1 CONTROL_{k,i,t}^1 + \varepsilon_{i,t}$$

$$CRASH_{i,t+1} = \beta_0^2 + \beta_1^2 Middle_{i,t} + \sum_k \gamma_k^2 CONTROL_{k,i,t}^2 + \varepsilon_{i,t+1}$$

where $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ are the measure of $CRASH_{i,t+1}$ of firm i in month $t + 1$, $CJR_{i,t}$ denotes the CJR measure, $Middle_{i,t}$ is the retail investor attention ($Abn\_SVI_{i,\,t}$), retail investor sentiment ($Abn\_Post_{i,\,t}$), and retail net-buys behavior ($SOIB_{i,\,t}$), $CONTROL_{k,i,t}$ is a set of control variables, including $NCSKEW_{i,t}$, $DUVOL_{i,t}$, $Ret_{i,t}$, $Sigma_{i,t}$, $LEV_{i,t}$, $BM_{i,t}$, $ROA_{i,t}$, $Size_{i,t}$, $Turn_{i,t}$, $Accm_{i,t}$ in Column (2) and Column (3) and $LEV_{i,t}$, $BM_{i,t}$, $ROA_{i,t}$, $Size_{i,t}$, $Turn_{i,t}$, $Accm_{i,t}$ in Column (1). Detailed definitions of all variables are shown in Appendix A.

Panel A: Retail investor attention			
Dependent variable =	$Abn\_SVI_{i,\,t}$ (1)	$NCSKEW_{i,\,t+1}$ (2)	$DUVOL_{i,\,t+1}$ (3)
CJR	0.085***		
	(5.748)		
Abn_SVI		0.020*	0.024**
		(1.875)	(2.327)
Constant	0.176***	0.417***	0.451***
	(6.496)	(19.404)	(15.022)
Controls	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.366	0.067	0.068

Panel B: Retail investor sentiment			
Dependent variable =	$Abn\_Post_{i,\,t}$ (1)	$NCSKEW_{i,\,t+1}$ (2)	$DUVOL_{i,\,t+1}$ (3)
CJR	0.145***		
	(7.582)		
Abn_Post		0.018***	0.022***
		(3.173)	(3.831)
Constant	0.070***	0.632***	0.589***
	(3.238)	(39.312)	(36.427)
Controls	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.292	0.103	0.106

Panel C: Retail net-buys behavior			
Dependent variable =	$SOIB_{i,\,t}$ (1)	$NCSKEW_{i,\,t+1}$ (2)	$DUVOL_{i,\,t+1}$ (3)
CJR	1.254**		
	(2.287)		
SOIB		0.090***	0.094***
		(3.154)	(3.294)
Constant	−0.288***	−1.700***	−1.891***
	(−7.437)	(−7.617)	(−8.213)
Controls	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.069	0.092	0.093

5.2. The influence of market mechanism

According to the transmission mechanism tests, we find that stock with positive information shocks attract more abnormal retail investor attention and aggravate their abnormal sentiment and net buys, driving the stock price to be overvalued. As time goes on, the information is gradually disclosed, and the stock price returns to its fundamental value. The higher the degree of stock overvaluation in the early stage, the greater the likelihood of price crashes in the later stage. In this section, we focus on two typical market mechanisms: (1) short sale constraints and (2) internet speech restrictions, to verify the importance of retail investors' behavior in this relationship between positive information shocks and crash risk.

Table 11

The effect of relaxing short-sale constraints. This table reports the subsample results of relaxing short-sale constraints while controlling for industry and month fixed effects. Column (1) and Column (3) employ the fixed effects regressions with stocks not relaxing the short-sale constraints, while Column (2) and Column (4) employ the fixed effects regressions with stocks relaxing the short-sale constraints. Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable =	$NCSKEW_{i,\ t+1}$		$DUVOL_{\ i,\ t+1}$	
	(1) No-relax	(2) relax	(3) No-relax	(4) relax
CJR	4.571***	3.570***	4.500***	3.798***
	(3.987)	(4.339)	(3.764)	(4.782)
NCSKEW	0.071***	0.021*		
	(3.221)	(1.669)		
DUVOL			0.061***	0.016
			(2.881)	(1.203)
Ret	2.794***	2.728***	2.908***	2.842***
	(7.623)	(8.551)	(8.436)	(8.802)
Sigma	1.501***	2.399***	1.597***	2.427***
	(6.387)	(15.683)	(7.086)	(14.435)
Lev	0.009***	−0.124***	0.009***	−0.124***
	(3.270)	(−3.817)	(2.948)	(−3.950)
BM	−0.059**	−0.004	−0.057**	−0.002
	(−2.383)	(−0.437)	(−2.147)	(−0.247)
ROA	0.615***	0.040	0.596**	0.017
	(2.935)	(0.920)	(2.564)	(0.462)
Size	0.047***	0.037***	0.048***	0.035***
	(2.990)	(3.041)	(2.905)	(2.741)
Turn	0.001*	0.001***	0.001*	0.001**
	(1.913)	(2.784)	(1.860)	(2.514)
Accm	0.001	−0.005	0.001	−0.007
	(1.125)	(−0.631)	(0.649)	(−1.010)
Constant	−2.010***	−1.104***	−2.124***	−1.138***
	(−5.584)	(−3.959)	(−5.712)	(−3.820)
Month fixed effects	YES	YES	YES	YES
Industry fixed effects	YES	YES	YES	YES
Adjusted R^2	0.086	0.085	0.083	0.083

5.2.1. Short sale constraints

The Chinese stock market had strict short sale constraints before August 2010. Under short sale constraints (high limits of arbitrage), pessimistic investor views cannot be effectively reflected in the stock prices, which means that the overvalued stock prices are not corrected in time (Miller, 1977; Hong and Stein, 2003; Lim, 2011; Chang et al., 2012), eventually leading to higher crash risk. To improve pricing efficiency, the Chinese stock market eased short sale constraints, and many stocks can be shorted after 2010. Besides, Deng et al. (2020) have pointed out that easing short sale constraints is conductive to lower crash risk. Therefore, we further explore the effect of short sale constraints on our main results.

According to the start and end time of the short selling of each stock, the stocks in the margin trading list are grouped. That is, the stock i in the short selling period is divided into the group of relaxing short sale constraints while the stock i before or after the short selling period is divided into the group of strict short sale constraints. Using these two groups, we re-estimate Eq. (4). The results are reported in Table 11.

The coefficients on $CJR_{i,t}$ in columns (2) and (4), which are the results of relaxing short sale constraints, are 3.570 and 3.798, respectively, which are significantly lower than those in columns (1) and (3) (4.571 and 4.500). This means that when the market relaxes short sale constraints, the positive impact of positive information shocks on the future crash risk has been alleviated, showing that the market mechanism has played the expected role, and indirectly implies the market demand for short selling trades.

5.2.2. Internet speech restrictions

The Interpretation was formally implemented on September 10, 2013. It mentioned that if defamatory and harmful posts published by citizens are viewed more than 5000 times or forwarded more than 500 times, citizens will face serious punishment, such as imprisonment of up to three years. The Interpretation adds cost and risk to investors publishing and spreading false opinions or information increase. The situation of spreading or maliciously guiding the abnormal sentiment of retail investors on Guba, forums, blogs, and other gathering places of retail investors may be effectively controlled. Therefore, we

Table 12

The effect of the implementation of internet speech restrictions policy. This table reports the subsample results of the implementation of Internet speech restrictions policy while controlling for industry and month fixed effects. Column (1) and Column (3) employ the fixed effects regressions before the implementation of the Interpretation, while Column (2) and Column (4) employ the fixed effects regressions after the implementation of the Interpretation. Detailed definitions of all variables are shown in Appendix A. The sample contains all A-share stocks listed on the Shanghai and Shenzhen Stock Exchange from January 2007 to December 2018. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable =	$NCSKEW_{i,\ t+1}$		$DUVOL_{\ i,\ t+1}$	
	(1) Before	(2) After	(3) Before	(4) After
CJR	4.015***	3.323***	4.060***	3.561***
	(3.542)	(4.928)	(3.706)	(5.506)
NCSKEW	0.020	−0.007		
	(1.532)	(−0.956)		
DUVOL			0.012	−0.015*
			(0.800)	(−1.753)
Ret	2.258***	1.389***	2.293***	1.390***
	(7.090)	(4.892)	(6.985)	(4.298)
Sigma	1.377***	2.301***	1.418***	2.303***
	(7.199)	(12.915)	(7.260)	(12.320)
Lev	0.002***	−0.068**	0.002***	−0.074***
	(2.879)	(−2.262)	(3.235)	(−2.600)
BM	−0.039***	0.008	−0.035**	0.010
	(−2.817)	(0.908)	(−2.427)	(1.170)
ROA	0.016	0.338*	0.009	0.200
	(1.136)	(1.646)	(0.682)	(0.994)
Size	0.030***	0.027**	0.027**	0.027*
	(2.817)	(1.963)	(2.438)	(1.862)
Turn	0.001***	0.001***	0.002***	0.001***
	(4.285)	(3.098)	(4.468)	(3.083)
Accm	0.001**	0.000	0.000	−0.001
	(2.071)	(0.031)	(0.665)	(−0.084)
Constant	−1.600***	−1.598***	−1.639***	−1.576***
	(−6.479)	(−5.303)	(−6.340)	(−4.872)
Month fixed effects	YES	YES	YES	YES
Industry fixed effects	YES	YES	YES	YES
Adjusted R^2	0.086	0.086	0.085	0.085

consider that the implementation of the Interpretation is likely to weaken the relationship between the positive information shocks and the crash risk by inhibiting the expression and spread of over-optimistic retail investor sentiment.

Based on the above analysis, taking September 2013 as the dividing line, we divide the total samples into two groups, and re-estimate Eq. (4), respectively Table 12. shows that there is a significantly positive correlation between cumulative positive jump returns and crash risk before and after the implementation of the Interpretation. However, comparing the coefficients, we can find that the coefficients of columns (2) and (4) (3.323 and 3.561) are significantly smaller than those of columns (1) and (3) (4.015 and 4.060), indicating the impact of positive information shocks on the crash risk has weakened after the implementation of the Interpretation. The evidence shows that the implementation of the Interpretation has a certain deterrent effect on users who spread false investment opinions and maliciously guide abnormal sentiment on the internet, thus effectively weakening the aggravating effect of the positive information shocks on the crash risk.

In sum, Tables 11 and 12 show that the specific market mechanism is conducive to weakening the positive impact of positive information shocks on crash risk. In addition, the results further confirm the transmission role of retail investors in this positive relationship.

6. Heterogeneity analysis

Our results have confirmed that positive information shocks increase the crash risk by affecting retail investor behavior. Some special firm characteristics and aggregate states tend to stimulate abnormal retail investor attention, abnormal sentiment, and net buying behavior (Avramov et al., 2016; Yao et al., 2019), and then affect the relationship between positive information shocks and crash risk. To deepen the understanding of the positive impact, we attempt to explore the potential differences in this relationship under different firm characteristics and aggregate states.

Table 13
The conditional analysis on firm characteristics. This table reports the regression results under different firm characteristics while controlling for industry and month fixed effects. The regression model is as follows:

$$CRASH_{i,t+1} = \beta_0 + \beta_1 CJR_{i,t} + \beta_2 CONDITION_{i,t} + \beta_3 CJR_{i,t} * CONDITION_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+1}$$

where $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ are the measure of $CRASH_{i,t+1}$ of firm *i* in month *t + 1*. Panel A presents results of $NCSKEW_{i,t+1}$ for the dependent variable. Panel B presents results of $DUVOL_{i,t+1}$ for the dependent variable. $CJR_{i,t}$ denotes the CJR measure, $CONDITION_{i,t}$ is a set of aggregate dummy variables, including firm's equity nature ($soe_{i,t}$), firm's size ($large_{i,t}$), and listing time ($age_{i,t}$). Detailed definitions of all variables are shown in Appendix A. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A Dependent variable =	$NCSKEW_{i,\ t+1}$		
	(1)	(2)	(3)
CJR	2.481***	3.199***	1.263***
	(4.883)	(6.790)	(2.731)
soe	−0.046**		
	(−2.398)		
*CJR *soe*	3.976***		
	(4.394)		
large		0.047***	
		(9.181)	
*CJR *large*		1.204*	
		(1.753)	
age			−0.001***
			(−3.276)
*CJR *age*			0.033***
			(4.902)
Constant	−0.293	−1.100***	−1.424***
	(−0.887)	(−18.481)	(−7.288)
Controls	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.100	0.057	0.084
Panel B Dependent variable =	$DUVOL_{i,\ t+1}$		
	(1)	(2)	(3)
CJR	2.645***	2.876***	1.451***
	(5.274)	(6.359)	(3.104)
soe	−0.044**		
	(−2.107)		
*CJR *soe*	3.867***		
	(3.991)		
large		0.043***	
		(8.660)	
*CJR *large*		1.406**	
		(2.121)	
age			−0.001***
			(−2.655)
*CJR *age*			0.033***
			(4.537)
Constant	−0.447	−1.059***	−1.478***
	(−1.292)	(−18.089)	(−7.226)
Controls	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.098	0.054	0.083

Based on the Eq. (4), we introduce the interaction term and establish Eq. (10) as follows:

$$CRASH_{i,t+1} = \beta_0 + \beta_1 CJR_{i,t} + \beta_2 CONDITION_{i,t} + \beta_3 CJR_{i,t} * CONDITION_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+1} \tag{10}$$

where $CONDITION_{i,t}$ are a series of conditional variables related to firm characteristics or aggregate states. $CJR_{i,t} * CONDITION_{i,t}$ is the interaction term between these conditional variables and cumulative positive jump returns, which aims

Table 14

The conditional analysis on different aggregate states. This table reports the regression results under different firm characteristics while controlling for industry and month fixed effects. The regression model is as follows:

$$CRASH_{i,t+1} = \beta_0 + \beta_1 CJR_{i,t} + \beta_2 CONDITION_{i,t} + \beta_3 CJR_{i,t} * CONDITION_{i,t} + \sum_k \gamma_k CONTROL_{k,i,t} + \varepsilon_{i,t+1}$$

where $NCSKEW_{i,t+1}$ and $DUVOL_{i,t+1}$ are the measure of $CRASH_{i,t+1}$ of firm i in month $t + 1$. Panel A presents results of $NCSKEW_{i,t+1}$ for the dependent variable. Panel B presents results of $DUVOL_{i,t+1}$ for the dependent variable. $CJR_{i,t}$ denotes the CJR measure, $CONDITION_{i,t}$ is a set of aggregate dummy variables, including up/down states (*up*), higher/lower sentiment (*senti*) and bear/bull market (*bull*). Detailed definitions of all variables are shown in Appendix A. Except for the CJR measure ($CJR_{i,t}$), all continuous variables are winsorized at the 1% level in each tail. The standard errors are corrected, using the double-clustering (firm and month) method, as discussed by Petersen (2009). *t*-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Panel A Dependent variable =	$NCSKEW_{i,t+1}$ (1)	(2)	(3)
CJR	2.192***	2.455***	1.673**
	(4.527)	(4.976)	(2.561)
up	−0.227***		
	(−6.067)		
*CJR *up*	1.780***		
	(3.242)		
senti		0.163***	
		(4.572)	
*CJR *senti*		1.394*	
		(1.920)	
bull			0.240***
			(4.598)
*CJR *bull*			1.744**
			(2.047)
Constant	−1.323***	−1.219***	−1.194***
	(−27.043)	(−20.879)	(−14.819)
Control	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.083	0.050	0.043

Panel B Dependent variable =	$DUVOL_{i,t+1}$ (1)	(2)	(3)
CJR	2.422***	2.191***	1.759**
	(4.895)	(4.631)	(2.454)
up	−0.375***		
	(−9.796)		
*CJR *up*	1.698***		
	(3.025)		
senti		0.151***	
		(4.258)	
*CJR *senti*		1.160*	
		(1.665)	
bull			0.394***
			(6.687)
*CJR *bull*			3.527***
			(3.727)
Constant	−1.234***	−1.212***	−1.657***
	(−24.662)	(−21.197)	(−18.712)
Control	YES	YES	YES
Month fixed effects	YES	YES	YES
Industry fixed effects	YES	YES	YES
Adjusted R^2	0.083	0.047	0.063

to capture the change of the effect of positive information shocks on crash risk caused by the change of firm characteristics or aggregate states.

6.1. Firm characteristics

Ownership structure, firm size, and firm listing time are selected. As a dummy variable, $soe_{i,t}$ equals 1 if firm i in month t is a state-owned enterprise. Following Yuan et al. (2016), $large_{i,t}$ equals 1 if the market value of firm i in month t is greater than the median market value of all firms in the same industry in the same month, and 0 otherwise. The third selected firm characteristic is listing time, and the variable $age_{i.t}$ is a monthly continuous variable to measure the listing time. For example, Ping An Bank (000,001. SZ) was listed in April 1991 and its listing time in January 2007 was 189 months Table 13. presents the results.

In line with the basic regression results, the coefficients on $CJR_{i,t}$ are positive at the 1% level in Table 10. The coefficients on $soe_{i.t}$ and $age_{i.t}$ are significantly negative, while the coefficient on $large_{i.t}$ is significantly positive, indicating that non-state-owned enterprises, firms with a large size, and young firms tend to face higher crash risk. Furthermore, the results of Table 10 also show that the interaction coefficients on three firm characteristics and cumulative positive jump returns are significantly positive, which means that the relationship between positive information shocks and crash risk is significantly different among stocks with different equity ownership, market value, and listing time. Specifically, the positive relationship is more prominent in SOEs, large firms, and firms that have been listed for a long time. Although the finding goes against our intuition, Hirshleifer (2001) once indicated that if mispricing is transmitted through social process, then some well-known stocks are most likely to be mispriced. Nofsinger (2001), and Kansal and Singh (2018) all provide consistent evidence to support this finding.

Therefore, we draw the conclusion that relatively speaking, stocks with such labels as "state-owned enterprise," "large scale" and "old brand" may be more likely to win the trust of retail investors in the event of positive information shocks, thus stimulating their abnormal attention, abnormal sentiment and speculative net buying behavior, causing stock price to deviate from its fundamental value, and eventually aggravating crash risk. The evidence further confirms the influence of retail investor behavior on the positive relationship.

6.2. Aggregate states

Baker and Wurgler (2006) document that aggregate states affect investor sentiment and behavior. Motivated by the above analysis, $up_{i,t}$ equals 1 when the market returns are higher than the risk-free return rate and 0 otherwise. Referring to Yi and Mao (2009), $senti_{i,t}$ equals 1 when CICSI, representing investor sentiment, is greater than the mean in the sample period, and 0 otherwise. Finally, referring to Pagan and Sossounov (2003), $bull_{i,t}$ equals 1 in a bull market, and 0 in a bear market Table 14. reports the results.

The coefficients on $CJR_{i,t}$ are positive at least at the 5% level. The interaction coefficients on three aggregate states and cumulative positive jump returns are significantly positive, which shows that the positive relationship is significantly different in different market ups and downs, market sentiment and market stages. In the rising stage, high market sentiment stage and bullish market trends, the positive relationship is stronger, supporting the prior literature.

7. Conclusion

Using large positive discontinuous changes in stock prices to proxy for positive information shocks, we explore the impact of positive information shocks on the crash risk. We find that the positive information shocks exacerbate future crash risk, and the short-term effect is stronger than the long-term effect. We also show that retail investors play an important role in this exacerbating effect. In particular, retail investors pay more attention to the stocks with positive price jumps, then their abnormal sentiment and net buying behavior push up the stock prices to deviate from their fundamental values, thus aggravating the crash risk in the future, which is consistent with the explanation of the crash risk by behavioral finance theory.

Our study contributes to the literature on information shocks and stock price responses (Chan, 2003; Tetlock, 2010; Kim et al., 2011b; Savor, 2012; Jiang and Zhu, 2017; Frank and Sanati, 2018), especially positive information shocks. Previous studies typically examine all news shocks together and thus get a sort of averaging of quite different effects (Chan, 2013; Savor, 2012; Tetlock, 2010, 2011). Our study contributes to this literature by showing that considering the sign of information shocks could potentially help reconcile the previous seemingly opposite findings (Frank and Sanati, 2018). Besides, we add to studies that focus on the trading mechanisms that underpin the price formation process. More importantly, our study also shows that considering the heterogeneity of investors rather than market aggregation process as a black box can provide important help for a more comprehensive understanding of the consequences of various market shocks.

Our study has important policy implications for regulators and policymakers. Since positive information shock may lead to investors' irrational behavior and aggravate the stock price crash in the future, market regulators should strengthen the monitoring of individual stocks with positive information shocks to prevent crash risk caused by excessive shock. Meanwhile, regulators also need to correctly guide and reasonably ease abnormal retail investor sentiment and trading behavior, preventing the stock price from plummeting after a short-term positive information shock and promoting the stable development of the capital market. More importantly, under situations of financial crises and high market uncertainty, the stock price will become extremely sensitive, especially to some unexpected information shocks, which can cause a large jump in the price and lead to more severe market consequences (Jiang and Kim, 2016). Finally, our research result also shows that relaxing the arbitrage restrictions such as short selling constraints and strengthening the regulation of information dissemination can positively reduce investors' irrational sentiment and market risk. Market regulators should thus clarify the corresponding policy effectiveness and strength the construction of capital market stability mechanism.

Declaration of Competing Interest

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgment

We would like to thank participants at the workshop of Tianjin University Financial Engineering Research Center, for valuable comments and suggestions. We acknowledge financial support from the National Natural Science Foundation of China [Grant No.: 72073101]. Ahmet Sensoy gratefully acknowledges support from the Turkish Academy of Sciences—Outstanding Young Scientists Award Program (TUBA-GEBIP).

Appendix A. variable definitions

Variables	Definitions
Stock price crash risk variables	
$NCSKEW_{i,t+h}$	The negative conditional skewness.
$DUVOL_{i,t+h}$	The down-to-up volatility.
Positive information shocks	
$CJR_{i,t}$	Cumulative positive jump returns, see Section 3.2. and Appendix B for detail
$CJR1_{i,t}$	Positive information shock estimated by rolling window.
$CJR2_{i,t}$	Using intraday jump return to measure positive information shock.
$CJR3_{i,t}$	Using large positive jump volatility to measure positive information shock.
Control variables	
$Ret_{i,t}$	Mean of firm-specific monthly returns.
$Sigma_{i,t}$	Standard deviation of firm-specific monthly returns.
$Lev_{i,t}$	Firm financial leverage.
$BM_{i,t}$	Book-to-market ratio.
$ROA_{i,t}$	Return on assets.
$Size_{i,t}$	The natural logarithm of the total assets.
$Turn_{i,t}$	Detrended average daily turnover.
$Accm_{i,t}$	Earnings management calculated by adjusted-Jones model.
$Liquidity_{i,t}$	Using the negative value of effective spread to measure liquidity.
$Volatility_{i,t}$	The average of daily realized volatility in month t.
$INST_{i,t}$	Institutional shareholding ratio. The value is the same in each quarter.
$Mgshare_{i,t}$	Management shareholdering ration. The value is the same in each quarter.
Intervening variables	
$Abn\_Post_{i,t}$	The monthly abnormal post of Guba Eastmoney Forum for stock *i* on month *t*, see Section 5.1.2. for details.
$Abn\_SVI_{i,t}$	The monthly abnormal SVI of stock *i* on month *t*, see Section 5.1.1. for details.
$SOIB_{i,t}$	The monthly net-buys of stock *i* on month *t*, see Section 5.1.3. for details.
Other variables	
$soe_{i,t}$	Equals 1 if the firm is SOE and to 0 otherwise.
$large_{i,t}$	Equals 1 where the circulation market value is higher than the median at the same industry and month *t* and to 0 otherwise.
$age_{i,t}$	The listing age, calculated by month.
$up_{i,t}$	Equals 1 when the market returns are higher than the risk-free return rate and to 0 otherwise.
$senti_{i,t}$	Equals 1 when CICSI is higher than the mean in the sample interval and to 0 otherwise.
$bull_{i,t}$	Equals 1 in the bull market and to 0 in the bear market.

Appendix B. jump test and sequential jump identification procedure

Let $\{S_{t_0},\ S_{t_1},\ \cdots,\ S_{t_N}\}$ be stock prices observed over the period $[0,\ T]$, where $t_0 = 0$, $t_N = T$ Jiang and Oomen (2008). show that

$$\mathrm{JS} = \frac{V_{(0,T)}N}{\sqrt{\Omega_{SWV}}}\left(1 - \frac{RV_N}{SWV_N}\right) \sim \mathrm{N}(0,\ 1)$$

where N is the number of observations sampled between zero and T. RV_N is realized variance, defined as

$$RV_N = \sum_{i=1}^{N} r_i^2$$

where $r_i = \ln(S_i/S_{i-1})$ is the continuously compounded logarithmic return. SWV_N is the variance swap, defined as

$$SWV_N = 2\sum_{i=1}^{N}(R_i - r_i) = 2\sum_{i=1}^{N}R_i - 2\ln\left(\frac{S_T}{S_0}\right)$$

where $R_i = S_i/S_{i-1} - 1$ is the simple return. Barndorff-Nielsen and Shephard (2006) show that BPV_N is a consistent estimator of $V_{(0,T)}$ and a consistent estimator of $V_{(0,T)}$ can be obtained based on the bi-power variation

$$BPV_N = \frac{1}{\mu_1^2}\sum_{i=1}^{N-1}|r_i||r_{i+1}|$$

where $\mu_p = 2^{p/2}\Gamma[\frac{p+1}{2}]/\sqrt{\pi}$, with $p = 6$. $\hat{\Omega}_{SWV}$ is a consistent estimator of Ω_{SWV}, defined as

$$\hat{\Omega}_{SWV} = \frac{1}{9}\mu_6\frac{N^3\mu_{6/p}^{-p}}{N-p+1}\sum_{i=0}^{N-p}\prod_{k=1}^{p}|r_{i+k}|^{6/p}$$

with $p = 6$.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Stat. 89 (4), 701–720.
Andersen, T.G., Bollerslev, T., Frederiksen, P., Ørregaard Nielsen, M., 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. J. Appl. Econ. 25 (2), 233–261.
Asness, C.S., Frazzini, A., Pedersen, L.H., 2019. Quality minus junk. Rev. Account. Stud. 24 (1), 34–112.
Avramov, D., Cheng, S., Hameed, A., 2016. Time-varying liquidity and momentum profits. J. Financ. Quant. Anal. 51 (6), 1897–1923.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Financ. 61 (4), 1645–1680.
Barber, B.M., Odean, T., 2007. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818.
Barber, B.M., Odean, T., Zhu, N., 2008. Do retail trades move markets? Rev. Financ. Stud. 22 (1), 151–186.
Barberis, N., Shleifer, A., Vishny, R., Zhao, C., 1998. A model of investor sentiment. Journal of financial economics 49 (3), 307–343.
Böhm, V., Chiarella, C., 2005. Mean variance preferences, expectations formation, and the dynamics of random asset prices. Math. Financ. Int. J. Math. Stat. Financ. Econ. 15 (1), 61–97.
Brennan, M.J., Huh, S.W., Subrahmanyam, A., 2016. Asymmetric effects of informed trading on the cost of equity capital. Manag. Sci. 62 (9), 2460–2480.
Brown, S., Hillegeist, S.A., Lo, K., 2009. The effect of earnings surprises on information asymmetry. J. Account. Econ. 47 (3), 208–225.
Brzeszczyński, J., Gajdka, J., Kutan, A.M., 2015. Investor response to public news, sentiment and institutional trading in emerging markets: a review. Int. Rev. Econ. Financ. 40, 338–352.
Campbell, J.Y., Kyle, A.S., 1993. Smart money, noise trading and stock price behaviour. Rev. Econ. Stud. 60 (1), 1–34.
Cao, H.jianH., and Ou-Yang, H. (2005). Bubbles and panics in a frictionless market with heterogeneous expectations. Working Paper, Available at SSRN 687488.
Chan, W.S., 2003. Stock price reaction to news and no-news: drift and reversal after headlines. J. Financ. Econ. 70 (2), 223–260.
Chang, E.C., Cheng, J.W., Pinegar, J.M., Yu, Y., 2012. Short-sale constraints: reductions in costs of capital or overvaluation? Evidence from Hong Kong. Pac. Basin Financ. J. 20 (3), 506–520.
Chang, E.C., Luo, Y., Ren, J., 2014. Short-selling, margin-trading, and price efficiency: evidence from the Chinese market. J. Bank Financ. 48, 411–424.
Cheema, M.A., Nartea, G.V., 2014. Momentum returns and information uncertainty: evidence from China. Pac. Basin Financ. J. 30, 173–188.
Chen, C., Kim, J.B., Yao, L., 2017. Earnings smoothing: does it exacerbate or constrain stock price crash risk? J. Corp. Financ. 42, 36–54.
Chen, X., Harford, J., Li, K., 2007. Monitoring: which institutions matter? J. Financ. Econ. 86 (2), 279–305.
Chen, Y., Xie, Y., You, H., Zhang, Y., 2018. Does crackdown on corruption reduce stock price crash risk? Evidence from China. J. Corp. Financ. 51, 125–141.
Chen, Z., Schmidt, A., 2021. Retail investor risk-seeking, attention, and the January effect. J. Behav. Exp. Financ., 100511.
Cheng, F., Chiao, C., Fang, Z., Wang, C., Yao, S., 2020. Raising short-term debt for long-term investment and stock price crash risk: evidence from China. Financ. Res. Lett. 33, 101200.
Cheng, F., Wang, C., Chiao, C., Yao, S., Fang, Z., 2021. Retail attention, retail trades, and stock price crash risk. Emerg. Mark. Rev., 100821.
Conrad, J., Cornell, B., Landsman, W.R., Rountree, B.R., 2006. How do analyst recommendations respond to major news? J. Financ. Quant. Anal. 41 (1), 25–49.
Conrad, J., Wahal, S., Xiang, R., 2015. High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics 116, 271–291.
Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. J. Financ. 66 (5), 1461–1499.
De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. J. Polit. Econ. 98 (4), 703–738.
DeBondt, W.F., Thaler, R., 1985. Does the stock market overreact? J. Financ. 40 (3), 793–805.
DeLisle, R.J., Ferguson, M.F., Kassa, H., Zaynutdinova, G.R., 2021. Hazard stocks and expected returns. J. Bank Financ. 125, 106094.
Deng, X., Gao, L., Kim, J.B., 2020. Short-sale constraints and stock price crash risk: causal evidence from a natural experiment. J. Corp. Financ. 60, 101498.
Deng, X., Hung, S., Qiao, Z., 2018. Mutual fund herding and stock price crashes. J. Bank Financ. 94, 166–184.

Ding, R., Hou, W., 2015. Retail investor attention and stock liquidity. J. Int. Financ. Mark. Inst. Money 37, 12–26.
Dionne, G., Zhou, X. (2019, November). Information Environments and High Price Impact Trades: Implication for Volatility and Price Efficiency. In Proceedings of Paris December 2019 Finance Meeting EUROFIDAI-ESSEC.
Engelberg, J.E., Reed, A.V., Ringgenberg, M.C., 2012. How are shorts informed? Short sellers, news, and information processing. J. Financ. Econ. 105 (2), 260–278.
Fang, Y., Hu, M., Yang, Q., 2018. Do executives benefit from shareholder disputes? Evidence from multiple large shareholders in Chinese listed firms. J. Corp. Financ. 51, 275–315.
Frank, M.Z., Sanati, A., 2018. How does the stock market absorb shocks? J. Financ. Econ. 129 (1), 136–153.
Ftiti, Z., Fatnassi, I., Tiwari, A.K., 2016. Neoclassical finance, behavioral finance and noise traders: assessment of gold–oil markets. Financ. Res. Lett. 17, 33–40.
Fu, J., Wu, X., Liu, Y., Chen, R., 2021. Firm-specific investor sentiment and stock price crash risk. Financ. Res. Lett. 38, 101442.
Gao, B., Yang, C., 2018. Investor trading behavior and sentiment in futures markets. Emerg. Mark. Financ. Trade 54 (3), 707–720.
Gao, Y., Wang, Y., Wang, C., Liu, C., 2018. Internet attention and information asymmetry: evidence from Qihoo 360 search data on the Chinese stock market. Phys. A 510, 802–811.
Govindaraj, S., Livnat, J., Savor, P.G., Zhao, C., 2013. Large price changes and subsequent returns. J. Invest. Manag. 12 (3), 31–58.
Han, X., Li, Y., 2017. Can investor sentiment be a momentum time-series predictor? Evidence from China. J. Empir. Financ. 42, 212–239.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Financ. 55 (3), 1263–1295.
Hirshleifer, D., 2001. Investor psychology and asset pricing. J. Financ. 56 (4), 1533–1597.
Hong, H., Stein, J.C., 2003. Differences of opinion, short-sales constraints, and market crashes. Rev. Financ. Stud. 16 (2), 487–525.
Hong, H., Lim, T., Stein, J.C., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. J. Financ. 55 (1), 265–295.
Hu, J., Li, S., Taboada, A.G., Zhang, F., 2020. Corporate board reforms around the world and stock price crash risk. J. Corp. Financ. 62, 101557.
Huang, Y., Qiu, H., Wu, Z., 2016. Local bias in investor attention: evidence from China's Internet stock message boards. J. Empir. Financ. 38, 338–354.
Jang, J., Kang, J., 2019. Probability of price crashes, rational speculative bubbles, and the cross-section of stock returns. J. Financ. Econ. 132 (1), 222–247.
Jarociński, M., Karadi, P., 2020. Deconstructing monetary policy surprises—the role of information shocks. Am. Econ. J. Macroecon. 12 (2), 1–43.
Jawadi, F., Louhichi, W., Ameur, H.B., Ftiti, Z., 2019. Do jumps and co-jumps improve volatility forecasting of oil and currency markets? Energy J. 40 (Special Issue).
Jawadi, F., Namouri, H., Ftiti, Z., 2018. An analysis of the effect of investor sentiment in a heterogeneous switching transition model for G7 stock markets. J. Econ. Dyn. Control 91, 469–484.
Jia, N., 2018. Corporate innovation strategy and stock price crash risk. J. Corp. Financ. 53, 155–173.
Jiang, G.J., Kim, W., 2016. Evaluating analysts' value: evidence from recommendation revisions around stock price jumps. Eur. J. Financ. 22 (3), 167–194.
Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise–a "swap variance" approach. J. Econ. 144 (2), 352–370.
Jiang, G.J., Yao, T., 2013. Stock price jumps and cross-sectional return predictability. J. Financ. Quant. Anal. 48 (5), 1519–1544.
Jiang, G.J., Zhu, K.X., 2017. Information shocks and short-term market underreaction. J. Financ. Econ. 124 (1), 43–64.
Jiang, G.J., Lo, I., Verdelhan, A., 2011. Information shocks, liquidity shocks, jumps, and price discovery: evidence from the US Treasury market. J. Financ. Quant. Anal. 527–551.
Jiang, G.J., Wei, S.X., and Zhu, K.X. (2018). "Jump-start" CAPM–information shocks and expected returns. Unpublished Working Paper.
Jin, L., Myers, S.C., 2006. R2 around the world: new theory and new tests. J. Financ. Econ. 79 (2), 257–292.
Kansal, P., Singh, S., 2018. Determinants of overconfidence bias in Indian stock market. Qual. Res. Financ. Mark. 10 (4), 381–394.
Kapadia, N., Zekhnini, M., 2019. Do idiosyncratic jumps matter? J. Financ. Econ. 131 (3), 666–692.
Kaplanski, G., Levy, H., 2010. Sentiment and stock prices: the case of aviation disasters. J. Financ. Econ. 95 (2), 174–201.
Kim, J.B., Li, Y., Zhang, L., 2011a. CFOs versus CEOs: equity incentives and crashes. J. Financ. Econ. 101 (3), 713–730.
Kim, J.B., Li, Y., Zhang, L., 2011b. Corporate tax avoidance and stock price crash risk: firm-level analysis. J. Financ. Econ. 100 (3), 639–662.
Kim, J.B., Wang, Z., Zhang, L., 2016. CEO overconfidence and stock price crash risk. Contemp. Account. Res. 33 (4), 1720–1749.
Kim, S.H., Kim, D., 2014. Investor sentiment from internet message postings and the predictability of stock returns. J. Econ. Behav. Organ. 107, 708–729.
Kubick, T.R., Lockhart, G.B., 2021. Industry tournament incentives and stock price crash risk. Financ. Manag. 50 (2), 345–369.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. J. Financ. 49 (5), 1541–1578.
Lee, C.M., Radhakrishna, B., 2000. Inferring investor behavior: evidence from TORQ data. J. Financ. Mark. 3 (2), 83–111.
Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: a new nonparametric test and jump dynamics. Rev. Financ. Stud. 21 (6), 2535–2563.
Li, W., Rhee, G., Wang, S.S., 2017. Differences in herding: individual vs. institutional investors. Pac. Basin Financ. J. 45, 174–185.
Lim, B.Y., 2011. Short-sale constraints and price bubbles. J. Bank Financ. 35 (9), 2443–2453.
Lin, T.C., Liu, X., 2017. Skewness, individual investor preference, and the cross-section of stock returns. Rev. Financ. 22 (5), 1841–1876.
Loewenstein, G., 2001. The creative destruction of decision research. J. Consum. Res. 28 (3), 499–505.
Ma, J., Xiong, X., Feng, X., 2021. News release and the role of different types of investors. Int. Rev. Financ. Anal. 73, 101643.
Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. J. Financ. 32 (4), 1151–1168.
Nartea, G.V., Kong, D., Wu, J., 2017. Do extreme returns matter in emerging markets? Evidence from the Chinese stock market. J. Bank Financ. 76, 189–197.
Ng, L., Wu, F., 2007. The trading behavior of institutions and individuals in Chinese equity markets. J. Bank Financ. 31 (9), 2695–2710.
Niu, Z., and Zhang, T. (2021). Stock returns on post macroeconomic announcement days. Available at SSRN 3495741.
Nofsinger, J.R., 2001. The impact of public information on investors. J. Bank Financ. 25 (7), 1339–1366.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108 (1), 1–28.
Pagan, A.R., Sossounov, K.A., 2003. A simple framework for analysing bull and bear markets. J. Appl. Econ. 18 (1), 23–46.
Park, T.J., Lee, Y., 2014. Informed trading before positive vs. negative earnings surprises. J. Bank Financ. 49, 228–241.
Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Rev. Financ. Stud. 22, 435–480.
Savor, P.G., 2012. Stock returns after major price shocks: the impact of information. J. Financ. Econ. 106 (3), 635–659.
Shleifer, A., Summers, L.H., 1990. The noise trader approach to finance. J. Econ. Perspect. 4 (2), 19–33.
Shleifer, A., Vishny, R.W., 1986. Large shareholders and corporate control. J. Polit. Econ. 94 (3), 461–488 Part 1.
Skinner, D.J., Sloan, R.G., 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. Rev. Account. Stud. 7 (2–3), 289–312.
Tetlock, P.C., 2010. Does public financial news resolve asymmetric information? Rev. Financ. Stud. 23 (9), 3520–3557.
Tetlock, P.C., 2011. All the news that's fit to reprint: do investors react to stale information? Rev. Financ. Stud. 24 (5), 1481–1512.
Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. J. Bank Financ. 41, 17–35.
Wang, A.W., Zhang, G., 2009. Institutional ownership and credit spreads: an information asymmetry perspective. J. Empir. Financ. 16 (4), 597–612.
Wang, M., Han, M., Huang, W., 2020. Debt and stock price crash risk in weak information environment. Financ. Res. Lett. 33, 101186.
Wen, F., Xu, L., Ouyang, G., Kou, G., 2019. Retail investor attention and stock price crash risk: evidence from China. Int. Rev. Financ. Anal. 65, 101376.
Xiao, Y., Yin, X., Zhao, J., 2020. Jumps, news, and subsequent return dynamics: an intraday study. J. Financ. Res. 43 (3), 705–731.
Xu, N., Jiang, X., Chan, K.C., Wu, S., 2017. Analyst herding and stock price crash risk: evidence from China. J. Int. Financ. Manag. Account. 28 (3), 308–348.
Xu, N., Li, X., Yuan, Q., Chan, K.C., 2014. Excess perks and stock price crash risk: evidence from China. J. Corp. Financ. 25, 419–434.
Yao, S., Wang, C., Cui, X., Fang, Z., 2019. Idiosyncratic skewness, gambling preference, and cross-section of stock returns: evidence from China. Pac. Basin Financ. J. 53, 464–483.

Yao, S., Wang, C., Fang, Z., Chiao, C., 2021. MAX is not the max under the interference of daily price limits: evidence from China. Int. Rev. Econ. Financ. 73, 348–369.
Yi, Z.G., Mao, N., 2009. Research on the measurement of investor sentiment in Chinese stock market: the CICSI's construction. J. Financ. Res. 11, 174–184 (in Chinese).
Yin, Y., Tian, R., 2017. Investor sentiment, financial report quality and stock price crash risk: role of short-sales constraints. Emerg. Mark. Financ. Trade 53 (3), 493–510.
Yu, B., Mizrach, B., Swanson, N.R., 2020. New evidence of the marginal predictive content of small and large jumps in the cross-section. Econometrics 8 (2), 19.
Yuan, R., Sun, J., Cao, F., 2016. Directors' and officers' liability insurance and stock price crash risk. J. Corp. Financ. 37, 173–192.
Zhang, J.D., Liao, W., Zhang, R.W., 2014. The effect of ordinary investor attention on volume and price of stock market: empirical evidence based on Baidu index. Account. Res. 08, 52–59 (in Chinese).