Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Information disclosure of COVID-19 specific medicine and stock price crash risk in China

Jiangjiao Duan*, Jingjing Lin

Business School, University of Shanghai for Science and Technology, Shanghai, China

ARTICLE INFO

JEL classification:
G1
G3

Keywords:
Stock price crash risk
Information disclosure
COVID-19

ABSTRACT

This study investigates the impact of information disclosure of COVID-19 specific medicine on stock price crash risk. To achieve this goal, crawling massive announcements of listed companies in China and using OCR (optical character recognition) to abstract text, texture analysis is adopted to count the keywords occurrence, which derives from a keywords list consisted of "specific medicine", "vaccine" and others. Based on information disclosure of COVID-19 specific medicine, empirical tests are conducted. The results show that the abnormal COVID-19 disclosure intensifies the degree of stock price crash risk. It proves that information asymmetry exactly exists in listed company information disclosure process in the securities market. It also provides new evidence that exaggerated information disclosure caused stock prices crash.

1. Introduction

After the Novel Coronavirus (COVID-19) outbreak, pharmaceutical research and development institutions in various countries begin to step up the research and development of vaccines. In this severe anti-epidemic situation, there is a phenomenon of a number of listed companies rubbing hot spots (Li, 2020). These companies claim to have successfully developed the raw material drug combination process and preparation technology of Remdesivir, and make false claims that the imitation of the specific drug can be completed in just ten days. After the company’s information is confirmed to be untrue, the company’s share price turns to a sharp decline.

This kind of extremely negative return on stock holdings is called stock price crash (Tian et al., 2018). In this field, most research focuses on managers’ hoarding negative news until a certain threshold triggers the stock price crash due to the concentrated outbreak of negative news. In accordance with agency theory, the hoarding of bad news is driven by agency problems and leads to information asymmetry and crash risk (Jin and Myers, 2006; Xu et al., 2013; Kong et al., 2021). Investors remember overconfident managers and companies with low social responsibility, so that they have foresight in their daily investment decision-making operations (Lee et al., 2019; Moradi et al., 2021). Some changes will suppress the stock price crash, including the emergence of new media increasing the difficulty for insiders to hide bad news (Ding et al., 2018; Sun et al., 2018; Zhang et al., 2018; Xie et al., 2016). Relatively new research is based on exaggerating positive news which triggers the stock price crash. From the perspective of motivation, listed companies tend to publish more positive information than negative information to affect stocks rate of return (Solomon, 2012). In addition, Zhao et al.

We appreciate the comments and suggestions of anonymous reviewers, editors and participants of the 3rd International Conference on Financial Regulation and Financial Stability.

* Corresponding author.
E-mail address: jjduan@ust.edu.cn (J. Duan).

https://doi.org/10.1016/j.frl.2022.102890
Received 7 February 2022; Received in revised form 24 March 2022; Accepted 13 April 2022
Available online 18 April 2022
1544-6123/© 2022 Elsevier Inc. All rights reserved.
find that the listed companies of related industries in China tend to exaggerate the positive news about ‘Internet Plus’, which rises the risk of stock price crash.

Related research on the influence of COVID-19 on capital market can be categorized into two aspects. One is about market crash risk (Corbet et al., 2022; Liu et al., 2021), and the other is about market efficiency (Heyden and Heyden 2021; Mazur et al., 2021). For example, Liu et al. (2021) investigate the impact of the COVID-19 pandemic on the stock market crash risk in China. Their study indicates that the pandemic increases stock market crash risk.

However, less literature has corresponding research on exaggerating positive information and COVID-19 specific medicine. Based on this, we use the special environment of COVID-19 to observe the information disclosure behavior of listed companies, and specifically the behavior of exaggerating the good news. This article selects January 2020 to December 2020 A-share data in China and text analysis data of company announcements issued on the exchange to examine the effect of COVID-19 specific on the stock price crash.

Our study contributes to the literature in several ways. First, this study supplements the research in the field about stock prices crash using COVID-19 information disclosure. By using python to crawl massive listed company announcement documents for text analysis. Second, from the perspective of exaggerating information, this study makes use of the special historical conditions of the COVID-19 epidemic. The sample documents to be analyzed include regular reports (annual, semiannual, quarterly reports) and other temporary announcements. In this paper, we find new evidence that exaggerated information caused stock prices crash.

The remainder of the paper is structured as follows. Section 2 is theoretical analysis and research hypothesis. Section 3 describes research design. Section 4 presents empirical results and robustness test. Section 5 concludes this paper.

2. Theoretical analysis and research hypothesis

This article mainly investigates whether information disclosure of COVID-19 specific medicine reduces or increases stock crash risk. Using the theory of information asymmetry, it is analyzed from the company level (Chen et al., 2001; Yuan, 2018). Information disclosure of listed companies is a bridge for companies to communicate with investors and the public. These include periodic reports and temporary announcements. After investors and the public obtain this information, they can be used as the main basis for investment decisions. However, due to information asymmetry or the existence of conflicts of interest, the company’s information disclosure may exaggerate positive information or hide negative information. Exaggerated disclosure of positive information includes not only false disclosures prohibited by regulation (Schrand and Walther, 2000), but also excessive publicity for legal compliance (Li, 2010; Solomon, 2012). For example, On February 13, 2020, WuChan ZhongDa (600704) publicly stated that they will soon develop COVID-19 specific medicine, although their main business has nothing to do with medicine. Later, it was verified by Shanghai Stock Exchange on March 1, 2020 that the announcement was untrue. Since then, the company’s share price turned to a sharp decline, as shown in Fig. 1. This behavior of WuChan ZhongDa (600704) is an exaggeration of the positive information disclosure of COVID-19 specific medicine.

During the outbreak of the COVID-19 epidemic, people are very concerned about special medicine, vaccines etc. related to this anti-epidemic, and so are investors. For example, Fig. 2 is the Baidu index for searching vaccines. From Fig. 2, we can see that the search volume for vaccines increased sharply from February 2020 and reached a peak on February 25, which the search volume exceeds 20,000.

In this situation, if a company discloses that it is related to the COVID-19 specific medicine, investors are very willing to buy the stock. However, once they realize that the news is exaggerated, it will have a negative impact on the company’s stock price. This
process is similar to a company hiding its negative news. Due to information asymmetry, the existence of the principal-agent problem, that is, the interests of shareholders and managers are not completely consistent, induces managers to pursue personal interests. Thereby the executives’ negative news hoarding will cause the stock price to fall (Jin and Myers, 2006; Kim et al., 2011; Piotroski et al., 2015; Fu et al., 2021). Similarly, due to information asymmetry, when exaggerated positive information disclosure exceeds the company’s threshold for its value, the company’s stock price will be higher than the company’s value (Zhao et al., 2020; Solomon, 2012). Once the stock price bubble has accumulated to a certain extent, it will have a negative impact on the company’s stock price and increase the risk of stock price crash.

From the previous research in this part, we put forward the first hypothesis of this article about information disclosure of COVID-19 specific medicine:

Hypothesis A: More disclosures, the higher the company’s stock price crash risk.

On the contrary, the causes of stock price crash can also be considered from the market level. According to the efficient market hypothesis put forward by the American financial economist Fama in 1965: If in a securities market, the price of securities completely reflects all available information, such a market is Efficient market. Harry Roberts distinguished three different levels of market efficiency for the first time: weak effective market, semi-strong effective market, and strong effective market. For some scholars, the stock price can reflect, to a large extent, the historical price information of listed companies in the stock market, as well, the public information of the company. Specifically, the company acclaims that it produces vaccine preparations by itself, which eliminates the information asymmetry between itself and investors, improves the company’s information transparency and better future prospect. In a word, we propose the second hypothesis:

Hypothesis B: More disclosures, the lower the company’s stock price crash risk.

3. Research design

3.1. The sample selection and data source

This article selects A-share listed companies in China from January 2020 to December 2020 as the research sample. The starting point for this article is 2020 year because of the official announcement about the new coronavirus epidemic. And the negative stock price bias finally generated annual data, meanwhile the keyword “vaccine” reflected by the Baidu search index peaks after 2020. The stock market transaction data comes from the CSMAR data center in China. In the process of selecting samples, financial listed companies are excluded, and listed companies with trading weeks less than 30 weeks are excluded. This article uses a 1% tailing process for continuous variables and obtain 3216 listed companies as sample. When we collect other data of control variables, four companies are delisted or sustain long-term trading suspension, whose announcements are unavailable. Finally, 3212 companies are the observations, and 384,429 announcements are downloaded to be analyzed.

3.2. Measuring information disclosure of COVID-19 specific medicine

In this paper, information disclosure of COVID-19 specific medicine is the core explanatory variable. The raw data of the number of information disclosures of COVID-19 specific medicine mainly comes from company announcements. These announcements include regular reports (annual, semiannual, quarterly reports) and other temporary announcements. The specific data acquisition and measurement steps are as follows: (1) The sample of all the announcements in the window period of those stock is crawled (by coding in python) from the Shanghai Stock Exchange and Shenzhen Stock Exchange. (2) Limit search etymology to exclude negative and interrogative sentences. Use the regulatory letter issued by the exchange to determine the inappropriate rhetoric in the disclosure of information violations as a key word. The key words mainly include: specific medicine, Remdesivir, antiviral, Gilead, mask, protective clothing, disinfectant, ventilator, vaccine, new coronavirus. (3) Text analysis of the announcements to exclude the announcement of the company’s production, revenue and other businesses affected by the COVID-19. (4) Coding in python to match keywords, the number of occurrences of keywords such as “specific medicine” in the announcement in the year is obtained, and ln_kwdis is measured.
by $\ln(1 + \text{thenumberofkeywordinoccurrence})$.

### 3.3. Measuring firm-specific crash risk

Following previous studies in the crash risk literature (Chen et al., 2001; Kim et al., 2014), this paper employs NCSKEW and DUVOL to measure stock price crash risk. Both measures are based on firm-specific weekly returns estimated as the residuals from the expanded market model. We first estimate the following expanded market model regression:

$$r_{it} = \beta_0 r_{m,t} + \beta_1 r_{m,t-1} + \beta_2 r_{m,t-2} + \beta_3 r_{m,t-3} + \beta_4 r_{m,t-4} + \epsilon_{it}$$  \hfill (1)

where $r_{it}$ represents the return on the $i^{th}$ stock in the $t^{th}$ week of a year. And $r_{m,t}$ denotes the market return which is the weighted sum of all stock circulation market value. The lead and lag items are included, and $\epsilon_{it}$ is the residual return from model as Eq. (1). And then, $W_{it} = \ln(1 + \epsilon_{it})$ is just the specific yield of the $i^{th}$ stock in the $t^{th}$ week.

Thus, for each firm $i$ over a fiscal-year period $t$, NCSKEW is calculated as Eq. (2):

$$\text{NCSKEW}_{it} = -\frac{\left[n(n-1)\sum_{\tau}W_{it}\right]}{\left[(n-1)(n-2)\sum_{\tau}W_{it}^2\right]^{3/2}}$$ \hfill (2)

where $n$ is the number of trading weeks for $i^{th}$ stock. Given that it multiplies by minus 1, the greater the negative stock price skewness, the higher the risk of stock price crash.

Another crash risk measure used in this paper, DUVOL is calculated as Eq. (3):

$$\text{DUVOL}_{it} = \ln\left\{\left[(n_u - 1)\sum_{d}R_u^2\right]/\left[(n_d - 1)\sum_{up}R_u^2\right]\right\}$$ \hfill (3)

For each firm $i$ over a fiscal-year period $t$, firm-specific weekly returns are separated into rising weeks and falling weeks by whether $W_{it}$ is greater than the annual average return. The number of rising and falling weeks are $n_u$,$n_d$. The standard deviation of the weekly returns of the two sub-samples is calculated to obtain $R_u$,$R_d$. The larger the DUVOL, the higher the risk of a stock price crash.

### 3.4. Empirical models

This paper constructs cross-sectional regression models Eq. (4) and Eq. (5) to investigate how specific medicine disclosure is associated with firm-specific stock price crash risk.

$$\text{NCSKEW}_{it} = \alpha + \beta \ln \text{kwd}\text{is}_{it} + \gamma \text{control}_{it} + \epsilon_{it}$$ \hfill (4)

$$\text{DUVOL}_{it} = \alpha + \beta \ln \text{kwd}\text{is}_{it} + \gamma \text{control}_{it} + \epsilon_{it}$$ \hfill (5)

where NCSKEW$_{it}$ is the negative conditional skewness of $i^{th}$ stock weekly returns over the fiscal year, and DUVOL$_{it}$ is the difference in stock price phase volatility. The explanatory variable $\ln \text{kwd}\text{is}_{it}$ is logarithm of the number of "specific medicine" occurrences of $i^{th}$ stock’s information disclosures. For the selection of control variables, we refer to previous studies (Chen et al., 2001; Kim et al., 2014; Xu et al., 2013; Chang et al., 2017; Peng et al., 2018). The control variable control includes the stock turnover rate (Dturn), stock yield (RE), market volatility (Sigma), company size (Csize), asset-liability ratio (Lev), company growth (MB), and return on assets (ROA). Detailed variable definitions are given in the Appendix A.

### Table 1

Descriptive statistics.

| Variables | Obs. | Mean   | Median | Standard Deviation |
|-----------|------|--------|--------|--------------------|
| NCSKEW    | 3212 | -0.4467| -0.4351| 1.0501             |
| DUVOL     | 3212 | -0.2009| -0.1931| 0.5957             |
| $\ln \text{kwd}\text{is}$ | 3212 | 1.6424 | 1.6094 | 1.1024             |
| Dturn     | 3212 | 0.3902 | 0.2054 | 0.8428             |
| RE        | 3212 | 0.2078 | 0.0453 | 0.6102             |
| Sigma     | 3212 | 0.0552 | 0.0524 | 0.0195             |
| Lev       | 3212 | 0.4370 | 0.4348 | 0.1968             |
| MB        | 3212 | 2.1263 | 1.6656 | 1.8707             |
| Csize     | 3212 | 22.4625| 22.2678| 1.3509             |
| ROA       | 3212 | 0.0341 | 0.2651 | 0.1000             |
4. Empirical results

4.1. Summary statistics

Table 1 shows all variables summary statistics. The mean value of explained variable NCSKEW is basically in line with most of papers in this field. The common range of the value is [−0.4, −0.2]. In Table 1, the mean of NCSKEW is −0.4467 and the median of it is −0.4351. The standard deviation of it is 1.0501, which indicates that there may be large differences in the risk of a stock crash within the sample. The mean of DUVOL is −0.2009, the median of it is −0.1931 and the standard deviation of it is 0.5957. The regression residuals used in the calculation of indicators NCSKEW and DUVOL are shown in the Appendix B. Every control variable is in an appropriate range refer to previous study in this field.

For all the A-share listed companies, there are 36% listed companies did not disclose any COVID-19 related information in their announcements. On the contrary, 64% listed companies have their announcements involves the COVID-19 related information. As can be seen from Table 2, the median of the occurrence is 4, which means more than half listed companies disclosed related information no less than 4. At the same time, the mean of the occurrence is 8.79, which means, in the whole stock market, the average keywords occurrence of the announcements with the keywords of “specific medicine” no less than 8. The maximum (298), the minimum (0) and the standard deviation (16.9741) implicate the distribution of “specific medicine” list keywords in the listed companies’ announcement is significant diverse in the A-share listed companies.

As Table 3 presents, the explained variables (NCSKEW, DUVOL) have pronounced association with, at least, half control variables at the 10% level of statistical significance. It justifies the selection of control variables. Moreover, the absolute value is no more than 0.5 consistent with the thumb rule. However, at a significance level of 10%, the coefficients of company size (Csize) and asset-liability ratio (Lev), market volatility (Sigma) and annual turnover rate (Dturn), market volatility (Sigma) and annual return on individual stocks (RE) are all close to 0.5. Further test with variance inflation factor is in Table 4.

As can be seen from Table 4, the variance inflation factor of all control variables is much less than the empirical value of 5. So, the control variables listed in Table 4 do not have multicollinearity. The selection of the set control variables of the model is based on the existing research. Multicollinearity test further guarantees the reliability of the results.

4.2. Regression analysis

4.2.1. “Specific medicine” disclosure and stock price crash risk

Table 5 presents the impact of information disclosure of specific medicine on crash risk. In Table 5, column (1) include control variables only, column (2) include all variables. The core explanatory variable ln_kwdis, which is the logarithm of the number of keywords disclosure, is significantly positively correlated with NCSKEW at 1% level. In addition, the model fit ($R^2$) is also improved with ln_kwdis. The regression results signify that more specific medicine keywords (keywords consisted in the dictionary) a listed company discloses in the announcement, the higher the company’s stock price crash risk.

Similarly, column (3) include control variables only, column (4) include all variables. Table 5 presents the regression results for another indicator to measure the risk of stock price crash, that is, the difference in volatility between stock price rise and fall (DUVOL). The correlation coefficient of ln_kwdis is also significantly positive. These regression results, as well, evidence that there is a significant positive correlation between the occurrence of keywords (keywords consisted in dictionary) and the risk of stock price crashes, confirming Hypothesis A again.

4.2.2. Controlling for industry fixed-effects

Given that there are many listed companies in A-shares that are vertically or even horizontally expanding, in other words, upstream and downstream subsidiaries or themselves invested some projects related to pharmaceuticals and medical preparations in their strategic horizontal expansion.

Thus, for the sake of integrality, the selection in the initial sample covers the entire A-share market in case. In fact, the shares of the pharmaceutical sector fluctuated intensively during the epidemic. The model is optimized with the fixed effects on the industry aspect. Specifically, the Eq. (4) and (5) are transformed as Eq. (6) and (7):

Table 2
The occurrence of the keywords in “specific medicine” list.

| Percentiles |        | Smallest |
|-------------|--------|----------|
| 1%          | 0      | 0        |
| 5%          | 0      | 0        |
| 10%         | 0      | 0        |
| 25%         | 1      | 0        |
| 50%         | 4      | 0        |
| 75%         | 10     | 214      |
| 90%         | 20     | 260      |
| 95%         | 30     | 274      |
| 99%         | 73     | 298      |

Obs: 3212, Sum of Wgt: 3212, Mean: 8.792653, Std. Dev.: 16.9741, Variance: 288.1202, Skewness: 7.681225, Kurtosis: 92.51023
\[ NCSKEW_i = a + \beta \ln \text{kwdis}_i + \gamma \text{control}_i + \delta \text{ind}_i + \varepsilon_i \]  
(6)

\[ DUVOL_i = a + \beta \ln \text{kwdis}_i + \gamma \text{control}_i + \delta \text{ind}_i + \varepsilon_i \]  
(7)

Here, when \( \text{ind}_i = 1 \), the listed company \( i \) affiliates with pharmaceutical sector. On the contrary, when \( \text{ind}_i = 0 \), it does not affiliate with pharmaceutical sector.

The regression results are reported in Table 6. From Table 6, we can see that the results still verify the Hypothesis A with industry fixed effects. The correlation coefficient of \( \ln \text{kwdis} \) is aligned with the previous result, which both pass the significance test level of 10%. That means, with the rise of “specific medicine”, “vaccine” and other related keywords occurrence in announcements of listed companies, the share crash risk is up.
4.3. Robustness test

4.3.1. Exaggerate positive information retest

In this paper, the occurrence of keywords is retested by four subsamples grouped by median and mean. Group that is not higher than the median (or lower than the mean) as a normal disclosure set, and a group that is higher than the median (or mean) as an exaggerated disclosure group.

Columns (1) to (4) of are subsample grouped by mean of $\ln kwdis$. As shown in Table 7, the control set is the normal disclosure set, while the experimental set is the excessive disclosure group. The results of column (2) and column (4) both implicate that exaggerated group of $\ln kwdis$ increase the share crash risk.

Similarly, columns (5) to (8) are subsample grouped by median of $\ln kwdis$. It evidences that, column (6) and column (8), for the exaggerated disclosure group, the correlation coefficient of $\ln kwdis$ is larger than that of the normal disclosure group. In general, the above regression results also further verify Hypothesis A. These results show that exaggerating the disclosure of keywords such as “vaccine” and “specific medicine” aggravates the risk of stock price crash. Two indicators (NCSKEW and DUVOL) that represent the share crash risk verified the Hypothesis A besides a robust regression result furthermore.

4.3.2. Self-selection effect test

In the analysis in Section 4.3.1, there may be crash risk due to firm heterogeneity, that is, self-selection effects. To address this concern, we employ the propensity-score matching (PSM) method. As in 4.3.1, the control set is the normal disclosure set, and the experimental set is the exaggerated disclosure set. By comparing the mean of $\ln kwdis$, a control set is selected according to whether it is exaggerated or not. During the PSM process, we choose to match firm variables including the stock turnover rate ($Dturn$), stock yield ($RE$), market volatility ($Sigma$), company size ($Csize$), asset-liability ratio ($Lev$), company growth ($MB$), and return on assets ($ROA$). According to the 1:1 nearest neighbor matching principle, a similar control sample is found for each experimental sample. Before re-estimating, a balance test is performed on the paired samples and the results are presented in Table 8. It can be seen from Table 8 that there is no significant difference between the experimental set and the control set in terms of company’s fundamental characteristics and earnings volatility, etc. The data using PSM captures the characteristic information of the experimental set to a large extent. Thus, the balance test supports the credibility of using PSM.

As shown in Table 9, after conducting the propensity score matching (PSM), we can see that the difference in stock price crash risk variables (NCSKEW, DUVOL) between experimental set and ctrl set becomes larger. And the t-statistic is significant after PSM, indicating that the crash risk difference is related to the exaggerated disclosure group.

We re-estimate the regression model using PSM paired samples. The results are reported in Table 10. From Table 10, we can see that the coefficient of $\ln kwdis$ in the experimental set is larger than that in the control set. Furthermore, regardless of whether NCSKEW or DUVOL is used, the coefficient of $\ln kwdis$ in the experimental group are significantly high. Combining Table 8 and Table 10, it can be seen that the differences in crash risk due to the differences in $\ln kwdis$ when the company’s fundamentals and characteristics such as volatility are very similar. These results are consistent with the previous analysis, again confirming that exaggerating the disclosure of keywords such as “vaccine” and “specific medicine” aggravates the risk of stock price crash.

Table 6
Regression with industry fixed-effects.

| Variables | (1) NCSKEW | (2) NCSKEW | (3) DUVOL | (4) DUVOL |
|-----------|------------|------------|-----------|-----------|
| ln_kwdis  | 0.0651**   | 0.0548*    | 0.480**   |           |
| Dturn     | -0.0538*   | -0.0548*   | 0.0422*   | -0.0429   |
| RE        | 0.178      | 0.394      | 0.489     | 0.187     |
| Sigma     | -17.39**   | -18.01**   | -14.11*** | -14.57*** |
| Lev       | 0.6955     | 0.9990     | 0.0225    | 0.0248    |
| MB        | 0.0336     | 0.0314     | 0.0181    | 0.0165    |
| Constant  | 2.323**    | 2.365**    | 1.723**   | 1.754**   |
| Observations | 3212       | 3212       | 3212      | 3212      |
| R-squared | 0.269      | 0.273      | 0.434     | 0.438     |

Note: Standard errors in parentheses. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.
5. Conclusion

Based on the actual COVID-19 information environment, this paper applies the stock price crash risk indicators used in quantities of previous studies, and examines the impact of keywords (such as specific medicine, Remdesivir, antiviral, Gilead, mask, protective clothing, disinfectant, ventilator, vaccine, new coronavirus) occurrence in the listed companies' announcements on stock price crash risk. The keywords counting result derives from 2020-year A-share listed companies' announcements in China.

Our empirical results show that: (1) The more keywords such as "specific medicine" appearing in the company’s announcement, the higher the company’s stock price crash risk. (2) Furthermore, with industry fixed-effects, the increase of the number of keyword disclosures aggravates the impact on the risk of stock price crash. (3) After grouping the full sample into a normal disclosure set and an exaggerated disclosure set, it finds that relative to listed companies that disclose keywords normally, the correlation between the risk of stock price crash and the number of keyword disclosures in the announcements of those listed companies with abnormal disclosures increased significantly. It directly evidences that during the epidemic period, there are indeed listed companies using COVID-19

### Table 7
Normal disclosure set & exaggerated disclosure set.

| Variables | By mean | Ctrl.set | Expm.set | By median | Ctrl.set | Expm.set | Ctrl.set | Expm.set |
|-----------|---------|----------|----------|-----------|----------|----------|----------|----------|
| ln kwdis  | 0.0337  | 0.107*** | 0.0431*  | 0.0757*** | 0.0274   | 0.101*** | 0.0421*  | 0.0745*** |
| Dturn     | −0.0714*** | −0.0322*** | −0.0575*** | −0.0224*** | −0.0699*** | −0.0333*** | −0.0566*** | −0.0228*** |
| RE        | −0.290**  | −0.394*** | −0.387*** | −0.665***  | −0.287**  | −0.597*** | −0.384*** | −0.669***  |
| Sigma     | −19.23*** | −16.79*** | −15.39*** | −13.57***  | −19.31*** | −16.76*** | −15.41*** | −13.58***  |
| Lev       | 0.0678  | 0.127    | −0.00626 | 0.0460    | 0.0636   | 0.135    | −0.00833 | 0.0511    |
| MB        | 0.0300  | 0.0428*** | 0.00866  | 0.0339*** | 0.0303   | 0.0437*** | 0.00912  | 0.0339***  |
| Fdist     | (0.0186) | (0.0111) | (0.0183) | (0.00846) | (0.0192) | (0.0108) | (0.0186) | (0.00833)  |
| Coze      | −0.126*** | −0.0476** | −0.0830*** | −0.0291**  | −0.129*** | −0.0464** | −0.0857*** | −0.0281**  |
| ROA       | −0.890*** | −0.309   | −0.713*** | −0.362***  | −0.810*** | −0.380**  | −0.662*** | −0.368**   |
| Cons.     | 3.314*** | 1.282*** | 2.399***  | 0.969***   | 3.396*** | 1.269***  | 2.456***  | 0.950***   |
| Obs.      | 16.48  | 1564     | 1648      | 1564      | 1606   | 1606     | 1606      | 1606      |
| R-sq.     | 0.1673  | 0.291    | 0.428     | 0.467     | 0.272   | 0.292    | 0.425     | 0.470      |

Note: Standard errors in parentheses. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

### Table 8
Balance test of the propensity-score matching (PSM) results.

| Variable | Ctrl.set | Expm.set | T-stat | p-value |
|----------|----------|----------|--------|---------|
| Dturn    | 0.190    | 0.175    | −0.230 | 0.817   |
| Sigma    | 0.055    | 0.054    | −1.330 | 0.185   |
| RE       | 0.196    | 0.197    | 0.040  | 0.965   |
| Lev      | 0.434    | 0.437    | 0.360  | 0.718   |
| MB       | 2.127    | 2.097    | −0.430 | 0.664   |
| Coze     | 22.471   | 22.501   | 0.580  | 0.562   |
| ROA      | 0.028    | 0.031    | 0.850  | 0.395   |

### Table 9
Differences in crash risk before and after the propensity-score matching PSM.

| Sample | NCSKEW | Ctrl.set | Expm.set | Difference | Std. Err. | T- stat |
|--------|--------|----------|----------|------------|-----------|---------|
| Before PSM | −0.454203518 | −0.498435193 | 0.044231674 | 0.035766744 | 1.24 |
| After PSM  | −0.388242932 | −0.504381974 | 0.116139042 | 0.038003236 | 3.06 |
| Sample   | Ctrl.set | Expm.set | Difference | Std. Err. | T- stat |
| Before PSM | −0.31488119 | −0.341382708 | 0.02650518 | 0.27232475 | 0.97 |
| After PSM  | −0.266579309 | −0.343560575 | 0.07692767 | 0.29801282 | 2.65 |

Based on the actual COVID-19 information environment, this paper applies the stock price crash risk indicators used in quantities of previous studies, and examines the impact of keywords (such as specific medicine, Remdesivir, antiviral, Gilead, mask, protective clothing, disinfectant, ventilator, vaccine, new coronavirus) occurrence in the listed companies’ announcements on stock price crash risk. The keywords counting result derives from 2020-year A-share listed companies’ announcements in China. Our empirical results show that: (1) The more keywords such as “specific medicine” appearing in the company’s announcement, the higher the company’s stock price crash risk. (2) Furthermore, with industry fixed-effects, the increase of the number of keyword disclosures aggravates the impact on the risk of stock price crash. (3) After grouping the full sample into a normal disclosure set and an exaggerated disclosure set, it finds that relative to listed companies that disclose keywords normally, the correlation between the risk of stock price crash and the number of keyword disclosures in the announcements of those listed companies with abnormal disclosures increased significantly. It directly evidences that during the epidemic period, there are indeed listed companies using COVID-19
information such as “specific medicine” for speculation, which might interfere with investor decision-making, leading to share fluctuations, rising the negative conditional skewness of $W_{i,t}$ (the share specific weekly return ratio) and phases volatility. It also provides new evidence that exaggerated information disclosure caused stock prices crash.

**CRediT authorship contribution statement**

**Jiangjiao Duan:** Supervision, Writing – review & editing. **Jingjing Lin:** Data curation, Software, Writing – review & editing.

**Declaration of Competing Interest**

No conflict of interest exits in the submission of this manuscript, and this manuscript is approved by all authors for publication.

**Acknowledgments**

This work was supported by the China National Social Science Foundation [21BGL087].

**Appendix A. Variable definitions**

| Variables  | Variables' definition |
|------------|-----------------------|
| NCSKEW     | the negative stock price skewness of the $i^{th}$ stock with the adjustment of market |
| DUVOL      | The difference in volatility between rising and falling phase of the $i^{th}$ stock |
| Dturn      | The stock turnover rate change equals to last-year one minus this-year one |
| ln_kwdis   | The logarithm of the occurrences of COVID-19 “specific medicine” |
| RE         | Individual stock annual return covering the cash bonus re-investment |
| Sigma      | The standard deviation of weekly return adjusted by market return |
| Lev        | Total debts divided by total assets |
| MB         | The market value of equity divided by the book value of equity |
| Csize      | The logarithm of the individual stock total assets |
| ROA        | The net profits divided by total assets |

**Appendix B. Regression of individual stock return on market return**

(continued on next page)
Appendix B is the result of regression Eq. (1). The residuals of this regression are basis used to calculate the two indicators of stock price crash risk (NCSKEW, DU VOL).

References

Chang, X., Chen, Y., Zolotoy, L., 2017. Stock liquidity and stock price crash risk. J. Financ. Quant. Anal. https://doi.org/10.1017/S00221090170000473.
Chen, J., Hong, H., Stein, J.C., 2001. Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. J. Financ. Econ. 61 (3), 345–381. https://doi.org/10.1016/S0378-4266(00)00076-4.
Corbet, S., Hou, Y.G., Hu, Y., Oxley, L., 2022. The influence of the COVID-19 pandemic on the hedging functionality of Chinese financial markets. Res. Inter. Bus. Financ. 59, 101510 https://doi.org/10.1016/j.ribaf.2021.101510.
Ding, H., Lv, C., Chen, Y., 2018. Investor information ability: disagreement of opinion and risk of stock price crash-evidence from 'shanghai stock exchange e interactive. Manag. World 34, 161–171. https://doi.org/10.17974/j.cnki.11-1235/f.2018.09.013.
Fu, J., Wu, X., Liu, Y., Chen, R., 2021. Firm-specific investor sentiment and stock price crash risk. Financ. Res. Lett. 38, 101442 https://doi.org/10.1016/j.frl.2020.101442.
Heyden, K.J., Heyden, T., 2021. Market reactions to the arrival and containment of COVID-19: an event study. Finance Res. Lett. 38, 101745 https://doi.org/10.1016/j.frl.2020.101690, 101690.
Jin, L., Myers, S.C., 2006. R² around the world: new theory and new tests. J. Financ. Econ. 79, 257–292. https://doi.org/10.1016/j.jfineco.2004.11.003.
Kim, J.B., Li, Y., Zhang, L., 2011. CFOs versus CEO: equity incentives and crashes. J. Financ. Econ. 101, 713–730. https://doi.org/10.1016/j.jfineco.2011.03.013.
Kim, Y., Li, H., Li, S., 2014. Corporate social responsibility and stock price crash risk. J. Bank. Financ. 43, 1–13. https://doi.org/10.1016/j.jbankfin.2014.02.013.
Kong, D., Shi, L., Zhang, F., 2021. Explain or conceal? Causal language intensity in annual report and stock price crash risk. Econ. Model. 94, 715–725. https://doi.org/10.1016/j.econmod.2020.02.013.
Lee, Y.C., Lu, Y.C., Wang, Y.C., 2019. Corporate social irresponsibility, CEO overconfidence, and stock price crash risk. Appl. Econ. Lett. 26, 1143–1147. https://doi.org/10.1080/13504851.2018.1540835.
Li, F., 2010. Textual analysis of corporate disclosures: a survey of the literature [J]. J. Account. Lit. 29 (2), 143–165 https://ssrn.com/abstract=1756926.
Liu, Y., 2020. Many Companies Boasted that they had “Special Effects” for new Crown Pneumonia and were Fined. China Economics Weekly, pp. 58–59. CNKI:SUN:JZXX.0.2020-05-026.
Liu, Z., Huynh, T.L.D., Dai, P.F., 2021. The impact of COVID-19 on the stock market crash risk in China. Res. Int. Bus. Financ. 57 https://doi.org/10.1016/j.ribaf.2021.101419.
Mazur, M., Dang, M., Vega, M., 2021. COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. Financ. Res. Lett. 38, 101690 https://doi.org/10.1016/j.frl.2021.101690.
Moradi, M., Appolloni, A., Zimon, G., Tarighi, H., Kamali, M., 2021. Macroeconomic factors and stock price crash risk: do managers withhold bad news in the crisis-ridden Iran market? Sustainability 13, 1–16. https://doi.org/10.3390/su13073688.
Peng, Y., Shen, J., Ni, X., 2018. Corporate “removing from reality to virtual” and financial market stability: based on the perspective of stock price collapse risk. Econ. Model. 53, 50–66.
Piotroski, J.D., Wong, T.J., Zhang, T., 2015. Political incentives to suppress negative information: evidence from Chinese listed firms. J. Account. Res. 53, 405–459. https://doi.org/10.1111/1475-679X.12071.
Schrand, C.M., Walther, B.R., 2000. Strategic benchmarks in earnings announcements: the selective disclosure of prior-period earnings components [J] Account. Rev. 75 (2), 151–177. https://doi.org/10.2308/acrrev.2000.75.2.151.
Solomon, D.H., 2012. Selective publicity and stock prices. J. Financ. 67 (2), 599–638. https://doi.org/10.1111/j.1540-6261.2012.01726.x, 2012.
Sun, Y., Fang, M., Yan, Q., 2018. Research on new media governance and stock price crash risk. Bus. Res. 135–145. https://doi.org/10.13902/j.cnki.613073688.
Tian, G., Shi, L., Qin, L., Yu, Z., 2018. Internet public opinion and its response to the information efficiency of listed companies. Syst. Eng. Theory Pract. 38, 46–66. https://doi.org/10.1201/b19788/2018030046-21.
Xie, D., Zheng, D., Cai, C., 2016. Is the pledge of controlling shareholders” equity a potential “landmine”? – A study based on the risk of stock price collapse. Manag. World 188, 128–140. https://doi.org/10.17974/j.cnki.11-1235/f.2016.05.011.
Xu, N., Yu, S., Yi, Z., 2013. Institutional investor herd behavior and stock price crash risk. Manag. World 31–43. https://doi.org/10.17974/j.cnki.11-1235/f.2013.07.004.
Yuan, Q., 2018. Summary of domestic and foreign literature on the risk of stock price collapse. Financ. Account. 124–129. https://doi.org/10.16144/j.cnki.issn1002-165 https://ssrn.com/abstract=3129276.
Zhang, J., Tang, X., Li, G., 2018. Can preventive supervision curb the risk of stock price collapse? – A study based on the inquiry letter of the exchange’s annual report. J. Manag. Sci. 21, 112–126. https://doi.org/10.13902/j.cnki.61307980.2018.010.009.
Zhao, C., Chen, S., Cao, W., 2020. “Internet Plus” information disclosure: substantive statement or strategic manipulation-evidence based on the risk of stock price crash. China Ind. Econ. 174–192. https://doi.org/10.19581/j.cnki.cjejournal.2020.03.020.