Finance Research Letters 38 (2021) 101838

Contents lists available at ScienceDirect

Finance Research Letters

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

The impact of COVID-19 on the Chinese stock market: Sentimental or substantial?

Yunchuan Sun *, Mengyuan Wu, Xiaoping Zeng, Zihan Peng

International Institute of Big Data in Finance, Business School, Beijing Normal University, Beijing 100875, China

A R T I C L E I N F O

Keywords:
COVID-19
Event study
Investor sentiment
Stock returns

A B S T R A C T

We investigate the impact of COVID-19 on Chinese stock market by an event study and examine the effect of individual investor sentiment on returns. The pandemic has an overall negative effect on stock market during the post-event window, which can't be explained by real losses. Results show a stronger positive correlation between individual investor sentiment and stock returns than usual. The impact on individual investor sentiment on stock returns is more significant for enterprises with high PB, PE and CMV, low net asset, and low institutional shareholding. Only 7 industries related to pharmacy, digitalization, and agriculture are boosted.

1. Introduction

The COVID-19 pandemic that started in early 2020 has led to a turbulence of financial market. The American stock market experienced circuit breakers twice in one week[1], and the cases in other countries were not much better. Most researchers have observed plummets during the pandemic, but the reasons remain unclear (Al-Awadhi et al., 2020; Fallahgoul, 2020; Nadeem Ashraf, 2020; Shehzad et al., 2020).

A rational explanation for the volatility would be substantial economic loss according to the efficient market hypothesis. If it holds, the region with more confirmed cases would suffer more substantial losses. Naturally, the profitability of companies in that area would be weakened, and their stock returns would decrease. From this perspective, as the center of the epidemic, the stock returns of companies in Hubei Province should be significantly lower than the average. This gap should continue to broaden as the situation worsens. In the same way, as the number of confirmed cases and the demand for medical supplies increase, the abnormal return rate of the pharmaceutical industry should also go significantly up correspondingly. However, our study shows that these are not the case. The stock returns of companies in Hubei show no differences with the market. The abnormal returns of pharmaceutical stocks did not last as well. This anomaly gives credence to the belief that stock market volatility during the COVID-19 epidemic cannot be explained simply by economic loss.

This paper explores the contribution of sentiment to stock market volatility during the epidemic by testing the following hypotheses. Two conditions should be met when major events affect stock returns through sentiment (Shan and Gong, 2012). First, the event leads to strong negative sentiment, such as panic and anxiety. Previous studies argued that public health hazards such as SARS

* Corresponding author.
E-mail addresses: yunch@bnu.edu.cn (Y. Sun), mengyuan_wu@mail.bnu.edu.cn (M. Wu), zengxp@mail.bnu.edu.cn (X. Zeng), 201711030433@mail.bnu.edu.cn (Z. Peng).

[1] The S&P 500 hit the 7% threshold decline on March 9 and March 12, halting trade during regular market hours for 15 minutes to ensure trading order. The last and only previous time of the "circuit breaker" was back in 1997.

https://doi.org/10.1016/j.frl.2020.101838
Received 23 September 2020; Received in revised form 26 October 2020; Accepted 3 November 2020
Available online 12 November 2020
1544-6123/

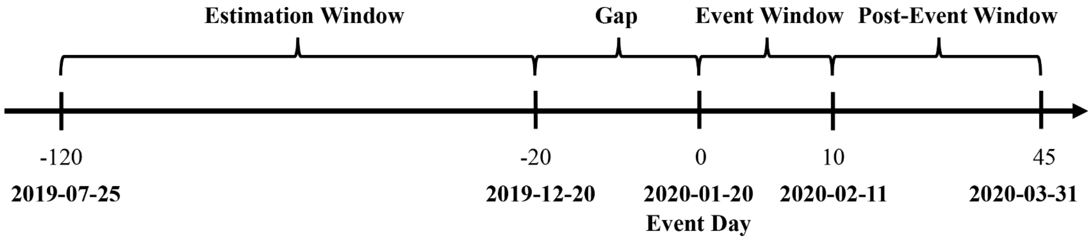

Fig. 1. Set up of the Event Study.

and Ebola can affect market sentiment (Tao, 2010). In the case of COVID-19, Liu et al. (2020) found that the virus outbreak had raised investors' fear of uncertainty. Baig et al. (2020) found that the overall sentiment declined during the pandemic. Second, the event causes lower yields on related stocks than usual. Donadelli et al. (2017) and Ichev and Marinč (2018) found that media coverage of pandemics had an impact on the stock prices of companies closer to the origin area and in the pharmaceutical industry. As for the pandemic this time, Baker et al. (2020) examined the US stock market and held that no previous infectious disease outbreak impacted the stock market as powerfully as COVID-19 did.

Event analysis is often used to measure the influence of major health emergencies. However, it fails to explore the contribution of different factors. Herein, we apply both event study and regression analysis in this study. First, we calculate the abnormal returns of the stock market during the pandemic and conduct a significance test. Then, we explore whether sentiment is explanatory to abnormal returns by regression. We also make an investigation on which kinds of stocks are more susceptible to sentiment during the pandemic. The main findings suggest that the epidemic has no significant effect on Chinese stock market except the pharmaceutical industry at first. As the epidemic spreads, significantly negative abnormal returns appear. The abnormal returns cannot be explained simply by real economic loss, so we turn to the spotlight on investor sentiment. Our results show that individual investor sentiment is positively correlated with stock market returns during the epidemic. Furthermore, stocks with high PB, PE, CMV, net asset and institutional shareholder ratios, as well as long listed years are more likely to be affected by the epidemic.

This study contributes to the literature in three ways. First, we contribute to the studies that have examined the stock market response to widespread disasters and provide empirical evidences for the sentiment effect in the Chinese stock market. Second, unlike most existing works that focus on overall market performance, we examine the effect of COVID-19 on individual investors, who are more susceptible to emotion and account for over 80% of the trading volume in the Chinese stock market.[2] Third, the sentiment is derived from analysis on big data of opinions text extracted from social platforms, in this way ensuring the sample size and credibility compared with investigations and market indicators.

2. Data and methodology

2.1. Data

Stock-related financial data are from the CSMAR database[3] covering the period from 25 July 2019 to 31 March 2020. A-share listed companies are selected as samples. In the panel data, we exclude samples with negative net assets in the annual report, with ST or PT designations,[4] within the financial sector and with missing values. The final sample number is 1914. The variables of enterprise features are processed by winsorizing at the level of 1% to avoid the influence of extreme values. The companies are then divided into 71 industries according to the China Securities Regulatory Commission.[5]

The sentiment data used in this work is GubaSenti established by International Institute of Big Data in Finance, BNU(http://ifind.bnu.edu.cn/), which captures the individual investor sentiment by text analytics on opinions from Guba – the biggest online financial social platform in China for individual investors to share and exchange their opinions and experiences on stocks (Sun et al., 2018; Sun et al., 2017).

2.2. Event study

Event study is applied in this work to identify abnormal returns in the stock market from the outbreak of COVID-19. January 20, 2020 is set as the event day when Nanshan Zhong, the senior expert on infectious disease in China, announced in a public interview that COVID-19 could be transmitted among people. It is the first time that human-to-human transmission has been confirmed possible officially. Existing literatures are not unanimous on the length of the estimation window. In this study, as shown in Fig. 1, we choose the 100 days before the event date as the estimation window. To observe the reversal effect, we also define the event window as 10

[2] Shanghai Stock Exchange Statistical Yearbook.
[3] https://www.gtarsc.com
[4] A stock receiving special treatment or delisting risk warning will be prefixed with "ST" or "PT". This prefix is used to indicate the risk of abnormal financial conditions or other abnormal conditions.
[5] http://www.csrc.gov.cn/pub/newsite/scb/ssgshyfljg/202004/t20200414_373793.html

Table 1
Summary Statistics.

Variable	Observations	Mean	SD	Min	Max
Panel A: Estimation Window					
Market return	100	0.001	0.008	-0.019	0.022
Stock return	324,729	0.001	0.022	-0.102	0.103
Sentiment	324,729	0.486	2.636	-19.880	21.463
Panel B: Event Window					
Market return	10	-0.005	0.030	-0.080	0.020
Stock return	31,475	-0.005	0.046	-0.103	0.104
Sentiment	31,475	0.431	2.728	-10.912	12.024
Panel C: Post-event Window					
Market return	36	-0.001	0.018	-0.039	0.034
Stock return	106,164	0.000	0.034	-0.109	0.103
Sentiment	106,164	0.627	2.550	-10.533	11.527

Notes: Table 1 reports summary statistics of the comprehensive A-share market daily return, sample stock daily return and investor sentiment. In panel A, the sample period is from July 25, 2019 to December 19, 2019. In panel B, the sample period is from January 20, 2020 to February 10, 2020. In panel C, the sample period is from February 11, 2020 to March 31, 2020. The market return and stock return data are derived from CSMAR database. The sentiment data is GubaSenti established by International Institute of Big Data in Finance, BNU(http://ifind.bnu.edu.cn/).

trading days after the event day and the post-event window as 10 days to 45 days after the event date. To avoid contaminated data, a 20-day gap is set between the event date and estimation window so that to capture the true abnormal returns. In the sensitive test, we estimate the 100, 50 and 150-day estimation window and gap respectively. The conclusions are robust.

The expected returns are derived using the Fama–French model. The ordinary least squares (OLS) regression is based on the following model:

$$R_{i,t} = \alpha + \gamma MKT_t + \delta SMB_t + \eta HML_t + \varepsilon_t \tag{1}$$

where $R_{i,t}$ represents the return of index i on date t in the estimation window, and MKT_t, SMB_t and HML_t are the three factors of the Fama–French model. Abnormal returns are calculated as follows:

$$AR_t = R_t - [\hat{\alpha} + \hat{\gamma} MKT_t + \hat{\delta} SMB_t + \hat{\eta} HML_t] \tag{2}$$

where R_t represents the actual return on date t in the event window. To measure the total impact of an event over a particular period (termed the "event window"), we add up individual abnormal returns to create a "cumulative abnormal return (CAR)" as follows:

$$CAR = \sum_{t=1}^{n} AR_t \tag{3}$$

Finally, after identifying all abnormal returns and cumulative abnormal returns during the event window, parametric tests are conducted to test significance. Evidence generally suggests that distributions of daily abnormal returns are relative to a normal distribution (Fama, 1965). According to it, our null hypothesis is CAR=0. If the epidemic has a significant positive impact on the stock price, the t-statistic should be significantly positive and vice versa.

In the robustness test, we change the length of the estimated window and the OLS regression model. The results remain consistent.

2.3. Panel regression model

Compared with an event study, a panel regression can better capture the time-varying relationship between dependent and independent variables due to its ability to extract changes from panel data and minimize estimation bias. Therefore, panel regression is used to control heterogeneity during estimation of the sentiment effect. The model is as follows:

$$R_{i,t} = \alpha_0 + \varphi MKT_t + \delta SMB_t + \eta HML_t + \gamma_1 SENT_{i,t} + \gamma_2 EW_t + \gamma_3 L_i + \varepsilon_{i,t} \tag{4}$$

where $R_{i,t}$ represents the return of individual stock i on day t and $SENT_{i,t}$ represents the sentiment index of individual stock i on day t. EW_t is a dummy variable representing the event window (0: the event window, 1: the post-event window). L_i is the dummy variable representing the region, and the values are 2, 1 and 0, which respectively indicate whether the company is located in Wuhan, other cities in Hubei and other regions in China. First, we run the regression using *SENT* as the independent variable alone and then gradually add *EW* and *L* to test the regional and reversal effects. To select between random or fixed effects, the Hausman test was implemented to verify the following hypothesis for each group. The use of the random or fixed effects model depended completely on the p-value. In the current study, the p-values are significant, so the null hypothesis is rejected, and we move on to the fixed effects model. To ensure the robustness of the results, we also conduct feasible generalized least squares (FGLS) estimation.

Considering that the investor sentiment and stock return may synchronously affect each other, endogeneity might exist and cause estimation bias. Therefore, the number of newly diagnosed sentiments in the previous day was taken as a proxy for investor sentiment

Table 2
Cumulative abnormal return for different event windows.

Event Windows	[0,9]		[10,45]	
Indices	Average CAR	T-Stats.	Average CAR	T-Stats.
Overall Market	0.007	4.010***	-0.016	-6.991***
Non-Pharmaceutical Industry	0.000	-0.062	-0.015	-6.305***
Pharmaceutical Industry	0.120	16.372***	-0.033	-3.441***
Hubei Province	0.009	0.729	-0.013	-0.880

Notes: Table 2 provides the event-study result. The event dates and windows are defined as in Fig. 1. *CAR* denotes mean cumulative abnormal returns. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 3
Cumulative abnormal return in different industries during event windows.

	Obs.	Average CAR	T-Stats.	Weight
Pharmaceutical Industry	197	0.120	16.372***	23.12%
Manufacture of computers, communication and other electronic equipment	293	0.033	4.879***	9.46%
Software and information technology services	177	0.044	5.159***	7.62%
Real estate	108	-0.052	-10.837***	5.49%
Manufacture of special purpose machinery	177	0.025	2.849***	4.33%
Manufacture of chemical raw materials and chemical products	203	0.015	2.098**	2.98%
Internet and related services	49	0.051	2.747***	2.44%
Business Service Industry	43	-0.057	-6.695***	2.40%
Other Industries	1919	-0.010	-5.002***	42.16%

Notes: Table 3 provides the event-study result in different industries during event windows. The event dates and windows are defined as in Fig. 1. *CAR* denotes cumulative abnormal returns. *Obs.* denotes the number of stocks in the industry. *Weight* measures the extent of the impact on the average CAR of the overall market and is calculated as below.

$$Weight = \frac{\text{The number of stocks in industry } i \times \text{ Average } CAR \text{ of industry } i}{\sum_{n=75}^{i} \text{The number of stocks in industry } i \times \text{ Average } CAR \text{ of industry } i}$$

***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively. In the event-study data, we do not exclude samples within the financial sector. Other Industries include all the 66 industries except the 9 industries listed above. See Appendix B for further details.

in the robustness test. The results show that though endogeneity exists, the estimation results are consistent.

3. Empirical findings

3.1. Event-study reports

Table 1 reports descriptive statistics. According to Panel B, stock returns and individual investor sentiment both react negatively after the event day. The standard deviation increases during the event window, which indicates that Chinese stock market yield decreases and volatility increases due to the epidemic. In addition to the event effect, the reversal effect is also observed. In the post-event window, both the return and investor sentiment rise, even exceeding the average level before the outbreak of COVID-19.

Table 2 reports the cumulative abnormal returns of the different event windows and market divisions. The results show that the cumulative abnormal return in the event window is positive, which means that the epidemic has a significant positive impact on the stock price in the short term. The second result, about pharmaceutical stocks, shows that the t-value is significantly positive, indicating the significant positive impact the epidemic has on the stock prices of pharmaceutical manufacturers. The event effect is insignificant for companies in Hubei.

These conclusions seem to be opposed to results of stock returns. Therefore, we explore the pharmaceutical industry and found that the cumulative abnormal return in pharmaceutical industry is far higher than the average and the sample size from pharmaceutical industry is large. As a result, its impact on the average CAR of the whole market reached 23.12%, leading to a bias in the overall results. Excluding the pharmaceutical industry, there was no significant cumulative abnormal return in the overall market during the event window. The results can be explained perfectly by the fact that the stocks in pharmaceutical industry were highly addressed by the investors during the epidemic. Similarly, due to quarantine, digitalization and information technology have become a magnet of investment, thus the abnormal returns of related industries are also significantly higher than other industries. In addition, during the event window, the epidemic is still at an early stage and its future is uncertain, so its impact on other industries is relatively limited.

In the post-event window, the cumulative abnormal return of the overall market decreases. The cumulative abnormal return of the overall market is significantly negative in this period, which means that the impact of the epidemic still exists in the long term and has a significant positive impact on the stock price. This phenomenon can be explained by the spread of the epidemic, which results in the extensive impact on the stock market and economy.

Table 4
Difference tests between industries and regions.

Event Windows	[-120, -21]		[0,9]		[10,45]	
Indices	CAR	P-Value	CAR	P-Value	CAR	P-Value
Pharmaceutical industry	0.024	-0.031***	0.120	-0.120***	-0.015	0.018*
Non-Pharmaceutical	-0.007		-0.000		-0.033	

Notes: Table 4 provides the results of tests of difference. The event dates and windows are defined as in Fig. 1. Pharmaceutical (Non-pharmaceutical) industry stocks are grouped according to the Industry Classification Scheme released by China Securities Regulatory Commission in 2020 Q1. *CAR* denotes mean cumulative abnormal returns. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 5
Panel data regression results of sentiment, reverse and region effect.

	Sentiment Effect		Reverse Effect		Region Effect	
	FE	FGLS	FE	FGLS	FE	FGLS
$SENT_i$	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***
	(114.55)	(113.85)	(114.79)	(114.07)	(114.79)	(114.07)
MKT	1.010***	1.011***	1.016***	1.017***	1.016***	1.017***
	(187.81)	(187.86)	(187.52)	(187.57)	(187.52)	(187.57)
SMB	0.604***	0.607***	0.625***	0.628***	0.625***	0.628***
	(54.62)	(54.86)	(55.40)	(55.65)	(55.40)	(55.65)
HML	-0.108***	-0.120***	-0.058***	-0.070***	-0.058***	-0.070***
	(-5.09)	(-5.63)	(-2.67)	(-3.18)	(-2.67)	(-3.18)
EW			-0.002***	-0.002***	-0.002***	-0.002***
			(-9.27)	(-9.31)	(-9.27)	(-9.31)
Location						
Other Hubei regions					0	-0.000
					(.)	(-0.29)
Wuhan City					0	-0.001
					(.)	(-0.77)
α_0	-0.003***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***
	(-25.43)	(-23.89)	(-4.74)	(-3.92)	(-4.74)	(-3.85)
adj. R-sq	0.477		0.477		0.477	
F	17404.1		13956.8		13956.8	
N	74455	74455	74455	74455	74455	74455

Notes: Table 5 provides the results from the estimation of the following regression specification.
$R_{i,t} = \alpha_0 + \varphi MKT_t + \delta SMB_t + \eta HML_t + \gamma_1 SENT_{i,t} + \gamma_2 EW_t + \gamma_3 L_i + \varepsilon_{i,t}$
Dependent variable $R_{i,t}$ represents the return of individual stock *i* on day *t* and $SENT_{i,t}$ represents the sentiment index of individual stock *i* on day *t*. EW_t is a dummy variable representing the event window(0: the event window, 1: the post-event window). L_i is the dummy variable representing the region, and the values are 2, 1 and 0, which respectively indicate whether the company is located in Wuhan, other cities in Hubei and other regions in China. First, we run the regression using *SENT* as the independent variable alone and then gradually add *EW* and *L* to test the regional and reversal effects. The adjusted R squares of the three fixed effect models are approximately equal. That is to say, the reverse and regional effect dummy variables have little effect on improving the fitting degree. Accordingly, the coefficients of *EW* and *L* are close to zero. T-stats are in parentheses below the coefficient estimates. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

We then test the differences in the cumulative abnormal return of stocks belonging and not belonging to firms in the pharmaceutical industry, and the registered places are in Hubei or other areas. The result shows that the abnormal return of pharmaceutical stocks is significantly higher than that of nonpharmaceutical stocks, while the cumulative abnormal returns of stocks registered in Hubei Province show no significant differences between those of other regions. One possible explanation is that the stocks of the pharmaceutical industry have some inherent characteristics that lead to excess returns that have never been discovered. However, the report in Table 4 shows that the difference between them is not as significant as it is before the pandemic, which supports the viewpoint that the industrial differences are caused by this epidemic and that the market is irrational.

3.2. Panel regression reports

The results in Table 5 prove that sentiment can significantly affect the overall market return during the epidemic. It also supports the position that the reversal effect is significant, indicating that stock returns depreciated during the post-event window.

There are three possible explanations for this phenomenon. First, as a result of the high infection rate and large volume of news concerning COVID-19, negative feelings are transmitted across China without distinction. Second, the Spring Festival is included in the event window, which is the transportation peak during the year. A large number of people traveled between Wuhan, Hubei Province and other Chinese cities during that time, which improved seriously the risk of infection and anxiety about the future. In addition, due to poor data availability, only companies registered in Hubei and Wuhan are included in the sample, thus excluding those who are not registered but conduct major business there.

Table 6
Panel data regression results of sentiment before and after the pandemic.

	[-100, -1] FE	[0, 45] FE	[-100, -1] FGLS	[0, 45] FGLS
$SENT_{i,t}$	0.003***	0.005***	0.002***	0.004***
	(-194)	(-114.55)	(-191.07)	(-113.85)
MKT	1.020***	1.010***	1.023***	1.011***
	(-199.69)	(-187.81)	(-199.81)	(-187.86)
SMB	0.733***	0.604***	0.738***	0.607***
	(-110.44)	(-54.62)	(-110.86)	(-54.86)
HML	-0.102***	-0.108***	-0.0996***	-0.120***
	(-11.94)	(-5.09)	(-11.67)	(-5.63)
α_0	-0.001***	-0.002***	-0.001***	-0.002***
	(-35.58)	(-25.43)	(-33.34)	(-23.89)
adj. R-sq	0.272	0.477		
F	30474.9	17404.1		
N	317901	74455	317901	74455
F	2988.71		754.78	
P-value	0.000***		0.000***	

Notes: Table 6 provides the results from the estimation of the following regression specification.
$R_{i,t} = \alpha_0 + \varphi MKT_t + \delta SMB_t + \eta HML_t + \gamma_1 SENT_{i,t} + \varepsilon_{i,t}$
Dependent variable $R_{i,t}$ represents the return of individual stock *i* on day *t* and $SENT_{i,t}$ represents the sentiment index of individual stock *i* on day *t*. In panel [-100, -1], the sample period is from August 22, 2019 to January 19, 2020. In panel [0, 45], the sample period from is from January 20, 2020 to to March 31, 2020. T-stats are in parentheses below the coefficient estimates. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively. The last two rows are the results of the coefficient difference test.

Table 7
Panel data regression results of stocks with different characteristics.

	FE $R_{i,t}$	**FGLS** $R_{i,t}$
$SENT_{i,t}$	0.003***	0.003***
	(23.66)	(25.30)
$SENT_{i,t} \times PE_i$	8.29E-06***	6.96E-06***
	(14.64)	(13.23)
$SENT_{i,t} \times PB_{i,t}$	3.06E-04***	3.08E-04***
	(13.84)	(15.27)
$SENT_{i,t} \times CMV_{i,t}$	2.84E-14***	1.12E-14***
	(8.44)	(3.97)
$SENT_{i,t} \times Year_i$	1.63E-05***	1.57E-05***
	(2.61)	(2.73)
$SENT_{i,t} \times ISR_i$	-1.40E-05	-1.55E-05*
	(-1.47)	(-1.78)
$SENT_{i,t} \times NA_i$	-3.22E-14***	-2.34E-14***
	(-6.31)	(-5.07)
MKT_t	1.013***	1.015***
	(174.33)	(174.33)
SMB_t	0.581***	0.585***
	(48.57)	(48.85)
HML_t	-0.111***	-0.123***
	(-4.84)	(-5.35)
ϕ_0	-0.003***	-0.002***
	(-23.68)	(-21.77)
adj. R-sq	0.487	
F	6023.2	
N	61832	61832

Notes: Table 7 provides the results from the estimation of the following regression specification.
$R_{i,t} = \varphi_0 + \varphi_1 SENT_{i,t} + \varphi_2 SENT_{i,t} \times Feature_i + \gamma MKT_t + \delta SMB_t + \eta HML_t + \varepsilon_t$
Dependent variable $R_{i,t}$ represents the return of individual stock *i* on day *t* and $SENT_{i,t}$ represents the sentiment index of individual stock *i* on day *t*. $Feature_i$ represents characteristics of stock *i*, including *PE, PB, CMV, Year, ISR* and *NA*. *PE* is the price earnings ratio calculated by daily stock market value and net income attributable to the parent company for the last four quarters. *PB* is the Price to book ratio calculated by daily stock market value and net asset published in the latest financial report. *CMV* is the current market value calculated by the daily number of tradable shares time the daily stock price. *ISR* denotes the institutional shareholding ratio. *Year* denotes the listed year by the event day January 20, 2020. *NA* denotes the amount of company net asset according to the latest annual report. The stock characteristics data are derived from CSMAR. T-stats are in parentheses below the coefficient estimates. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 8
Panel data regression results of stocks in different industries.

	FE	FGLS
	$R_{i,t}$	$R_{i,t}$
$SENT_{i,t}$	0.005***	0.004***
	(8.09)	(7.96)
$SENT_{i,t} \times Industry_i$		
Oil and gas extraction	-0.004***	-0.004***
	(-3.23)	(-3.25)
Petroleum and nuclear power processing	-0.004***	-0.003**
	(-2.76)	(-2.42)
Road transport	-0.004***	-0.003***
	(-4.79)	(-4.42)
Railway transport industry	-0.004**	-0.003*
	(-2.00)	(-1.84)
Coal mining and washing	-0.006***	-0.003***
	(-4.43)	(-4.04)
MKT_t	1.014***	1.015***
	(184.07)	(184.16)
SMB_t	0.566***	0.571***
	(49.99)	(50.35)
HML_t	-0.068***	-0.080***
	(-3.11)	(-3.69)
ϕ_0	-0.003***	-0.003***
	(-25.22)	(-23.46)
adj. R-sq	0.488	
F	888.0	
N	68083	68083

Notes: Table 8 provides the results from the estimation of the following regression specification.
$R_{i,t} = \varphi_0 + \varphi_1 SENT_{i,t} + \varphi_2 SENT_{i,t} \times Industry_i + \gamma MKT_t + \delta SMB_t + \eta HML_t + \varepsilon_t$
Dependent variable $R_{i,t}$ represents the return of individual stock i on day t and $SENT_{i,t}$ represents the sentiment index of individual stock i on day t. $Industry_i$ is the dummy variable based on the Industry Classification Scheme released by China Securities Regulatory Commission in 2020 Q1. Integrated industry is used as the reference group. Companies of the integrated industry have similar revenue and profit in many industries, and no obvious development planning and share holding background. T-stats are in parentheses below the coefficient estimates. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively. See Appendix C for further details.

To explore the special impact of the epidemic on investor sentiment, the data within the estimation window for regression were used. Table 6 shows that sentiment both before and after the event day has a significant impact on the return of individual stocks. We further conduct the Chows test to examine the coefficient difference. For FE and FGLS, the P values are both 0.000. The results indicate that the impact of sentiment during the epidemic has increased significantly from its usual level.

3.3. Characteristic effect

We set up an interactive variable in model (5) to study the performance of stocks with different characteristics during the epidemic.

$$R_{i,t} = \varphi_0 + \varphi_1 SENT_{i,t} + \varphi_2 SENT_{i,t} \times Feature_i + \gamma MKT_t + \delta SMB_t + \eta HML_t + \varepsilon_t \tag{5}$$

In the above, $Feature_i$ represents characteristics of stock i. Table 7 indicates that stocks with higher P/E ratio, P/B ratio, and circulating market value, and lower institutional shareholding ratio and net asset, are more susceptible to sentiment during the epidemic. This is possibly because the more resilient a company is to unexpected risks, the less anxious its investors are.

3.4. Industry effect

During the epidemic, different industries have experienced different turbulences (Baek et al., 2020; Mazur et al., 2020). From our observation, due to the significant increase in demand for pharmaceutical products, most of the top 10 stocks with a cumulative excess return during the event window are related to the pharmaceutical industry. In the post-event window, the food industry has become a new focus due to the food shortage alert released by the UN. Therefore, the role of industry effects in epidemics is worth exploring.

We take the comprehensive class in the industry category as the benchmark group, thereby adding the industry dummy variable to model 4. The results show that nearly half of the industries are significantly affected by sentiment, as 37 of the 71 dummy variables are significant. Among them, 30 industries have a weakening effect on the positive impact of sentiment. The five industries with the most negative impact are listed in Table 8. The reason may lie in the oversupply of oil and the economic downturn, as well as the restrictions on travel. The sentiment effect of only 7 industries—the Pharmaceutical Industry, Internet and Related Services, Processing of Food from Agricultural Products, Software and Information Technology Services, Manufacturing of Computers, Communication and Other Electronic Equipment, Farming and Education—is boosted. We can find these industries are related to pharmacy, digitalization including online education, and food crises. Only parts of the regression results are shown in Table 8 (see more in Appendix C).

4. Conclusion

This paper evaluates the influence of COVID-19 on China's stock market based on an event study and panel regression. This research adds to the literature, as it explores the unexpected outbreak effects on Chinese financial markets of a feared disease. We further prove that pandemics can cause widespread negative sentiment, thus leading to investor anxiety and market turbulence. The volatility of stock returns during the epidemic is influenced by sentiment and can't be explained solely by economic losses. Stocks with different financial characteristics and in different industries are affected differently. Among the thirty seven industries with significant industry effects, the sentiment effects of only seven industries, related to the Internet, education, medical manufacturing and agricultural production, are boosted. The five industries in which the positive sentiment effects have been significantly weakened are all related to oil, fuel and transportation.

Due to the high unpredictability of epidemics, investors could get excess return by holding bellwether stocks of the pharmaceutical industry in the first stage. Then, investors should gradually reduce stockholdings in the pharmaceutical industry and increase stockholdings highlighted by government. In addition, stocks with high risk factors, such as high P/E and P/B ratios, high CMV, low institutional shareholding ratio, and low net assets, should be avoided during the middle and late stages of the epidemic.

CRediT authorship contribution statement

Yunchuan Sun: Supervision, Writing - review & editing, Project administration, Resources, Conceptualization, Methodology, Software, Data curation, Writing - original draft. **Mengyuan Wu:** Conceptualization, Software, Validation, Visualization, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Xiaoping Zeng:** Supervision, Writing - review & editing. **Zihan Peng:** Writing - original draft, Validation, Investigation.

Acknowledgements

We thank all members of the International Institute of Big Data in Finance, Business School, BNU for the sentiment data and fruitful discussions.

Appendix A. Nomenclature

Nomenclature
$SENT_{i,t}$: Individual investor sentiment index of stock *i* on day *t*.
$SENT_{m,t}$: Individual investor sentiment index of the overall market on day *t*.
R_i: The daily return of stock *i* considering cash dividends reinvested.
Nnind: The industry code announced by China Securities Regulatory Commission.
PB: Price to book ratio. Calculated by daily stock market value and net asset published in the latest periodic report.
PE: Price earnings ratio. Calculated by daily stock market value and net income attributable to the parent company for the last four quarters.
CMV: Current market value. Calculated by the daily number of tradable shares time the daily stock price.
ISR: Institutional shareholding ratio. Quarterly data.
Year: The listed year by the event day.
NA: Net asset according to the latest annual report.
Regplc: Company registration place dummy variables. Wuhan =2, Hubei =1, other places =0.
EW: Event window dummy variables. Event window [0,10)=0, post-event window [10,46)=1.

Appendix B. Cumulative abnormal return in different industries during event windows

Table 9.

Table 9
Cumulative abnormal return in different industries during event windows

	Obs.	CAR	T-value	Weight
Pharmaceutical Industry	197	0.12	16.372***	23.12%
Manufacture of computers, communication and other electronic equipment	293	0.033	4.879***	9.46%
Software and information technology services	177	0.044	5.159***	7.62%
Real estate	108	-0.052	-10.837***	5.49%
Manufacture of special purpose machinery	177	0.025	2.849***	4.33%
Manufacture of chemical raw materials and chemical products	203	0.015	2.098**	2.98%
Internet and related services	49	0.051	2.747***	2.44%
Business Service Industry	43	-0.057	-6.695***	2.40%
Manufacture of alcohol, beverages, and refined tea	34	-0.064	-7.331***	2.13%
Production and distribution of electric power and heat power	66	-0.028	-5.059***	1.81%
Water transport	26	-0.062	-5.282***	1.58%
Civil engineering	57	-0.026	-3.791***	1.45%
Manufacture of non-metallic mineral products	77	-0.019	-2.303**	1.43%
Capital markets services	41	-0.035	-6.657***	1.40%
Manufacture of general-purpose machinery	109	-0.013	-1.384	1.39%
Ecological protection and environmental management	35	0.04	2.601**	1.37%
Decoration	23	-0.058	-3.716***	1.30%
Road transport	33	-0.04	-4.196***	1.29%
Manufacture of chemical fibers	20	0.06	2.026*	1.17%
Manufacture of textiles	30	0.038	1.978*	1.12%
Retail trade	76	-0.015	-1.368	1.12%
Management of public facilities	15	-0.075	-5.492***	1.10%
Mining and washing of coal	24	-0.046	-8.944***	1.08%
Manufacture of furniture	19	-0.054	-3.202***	1.00%
Wholesale trade	70	0.014	1.173	0.96%
Smelting and processing of non-ferrous metals	61	-0.016	-1.365	0.95%
Processing of food from agric. products	44	0.022	2.063**	0.95%
Manufacture of railway, ships, aerospace and other transportation equipment	41	-0.022	-1.51	0.88%
Manufacture of metal products	53	-0.015	-1.45	0.78%
Other financial activities	15	-0.051	-3.361***	0.75%
Air transport	11	-0.066	-4.659***	0.71%
Production and distribution of gas	18	-0.04	-3.428***	0.70%
Manufacture of articles for culture, education, art, sports, and entertainment	12	0.058	1.562	0.68%
Manufacture of automobiles	114	-0.006	-0.575	0.67%
Postal services	4	0.169	3.856**	0.66%
Ancillary mining activities	14	-0.044	-2.033*	0.60%
Manufacture of foods	38	0.015	1.217	0.56%
Smelting and processing of ferrous metals	28	-0.02	-1.54	0.55%
News and publishing	22	0.025	1.431	0.54%
Manufacture of leather, fur, feather and related products; footwear industry	11	-0.048	-1.796	0.52%
Other manufacturing	13	-0.038	-2.009*	0.48%
Manufacture of textiles, clothing; apparel industry	32	-0.015	-0.716	0.47%
Processing of timber, manufacture of wood, bamboo, rattan, palm, and straw products	7	-0.068	-3.292**	0.47%
Comprehensive use of waste resources	3	0.157	5.038**	0.46%
Farming	14	0.033	1.924*	0.45%
Monetary and financial services	25	-0.018	-4.92***	0.44%
Research and experimental development	4	0.112	4.076**	0.44%
Accommodation	6	-0.074	-5.503***	0.43%
Production and distribution of tap water	14	0.03	1.311	0.41%
Complex	15	0.028	1.012	0.41%
Printing and recorded media	11	0.034	0.999	0.37%
Healthcare	12	0.029	1.112	0.34%
Manufacture of rubber and plastics	65	-0.005	-0.507	0.32%
Extraction of petroleum and natural gas	6	-0.054	-3.542**	0.32%
Professional technical service industry	42	-0.007	-0.591	0.29%
Education	8	0.035	0.888	0.27%
Insurance	6	-0.04	-3.017**	0.23%
Storage	7	-0.034	-1.042	0.23%
Railway transport	4	-0.057	-4.621**	0.22%
Telecommunications, radio, television, and satellite transmission services	14	-0.016	-1.395	0.22%
Manufacture of measuring instruments	42	-0.005	-0.413	0.21%
Manufacture of electrical machinery and equipment	203	-0.001	-0.175	0.20%
Animal husbandry	12	0.015	0.717	0.18%
Catering	2	-0.086	-7.094*	0.17%
Leasing	3	-0.057	-6.141**	0.17%
Mining and processing of non-ferrous metal ores	21	-0.007	-0.389	0.14%
Construction of buildings	2	-0.064	-9.806*	0.13%
Mining and processing of ferrous metal ores	5	-0.022	-0.617	0.11%

(continued on next page)

Table 9 (*continued*)

	Obs.	CAR	T-value	Weight
Loading/unloading, removal, and other transport services	2	-0.053	-2.574	0.10%
Radio, film, television, and (other) audio-visual media	19	-0.005	-0.147	0.09%
Processing of petroleum, coking, processing of nuclear fuel	14	0.006	0.141	0.08%
Manufacture of paper and paper prod.	24	0.003	0.132	0.07%
Culture and arts	8	-0.003	-0.049	0.02%
Forestry	2	-0.008	-0.544	0.02%
Fisheries	7	0.002	0.177	0.01%

Notes: This table provides the event-study result in different industries during event windows. The event dates and windows are defined as in Fig. 1. *CAR* denotes mean cumulative abnormal returns. Obs. denotes the number of stocks in the industry. Weight measures the extent of the impact on the average CAR of the overall market and is calculated as below.

$$Weight = \frac{\text{The number of stocks in industry } i \times \text{Average } CAR \text{ of industry } i}{\sum_{n=75}^{i} \text{The number of stocks in industry } i \times \text{Average } CAR \text{ of industry } i}$$

***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Appendix C. Panel data regression results of stocks in different industries

Table 10

Table 10
Panel data regression results of stocks in different industries

	FE	FGLS
	$R_{i,t}$	$R_{i,t}$
$SENT_{i,t}$	0.00483***	0.00419***
	(8.09)	(7.96)
$SENT_{i,t} \times Industry_i$		
Farming	0.002*	0.002***
	(1.82)	(2.80)
Forestry	-0.001	-0.001
	(-0.66)	(-0.31)
Animal husbandry	0.001	0.001
	(1.47)	(1.47)
Fisheries	-0.003*	-0.002
	(-1.79)	(-1.27)
Mining and washing of coal	-0.004***	-0.003***
	(-4.43)	(-4.04)
Extraction of petroleum and natural gas	-0.004***	-0.004***
	(-3.23)	(-3.25)
Mining and processing of ferrous metal ores	-0.003	-0.003
	(-1.47)	(-1.37)
Mining and processing of non-ferrous metal ores	-0.000	0.000
	(-0.56)	(0.17)
Ancillary mining activities	-0.003***	-0.002***
	(-3.20)	(-2.83)
Processing of food from agric. products	0.001**	0.002***
	(2.06)	(3.50)
Manufacture of foods	0.000	0.000
	(0.50)	(0.45)
Manufacture of alcohol, beverages, and refined tea	-0.003***	-0.002***
	(-3.51)	(-3.18)
Manufacture of textiles	-0.001*	-0.001
	(-1.93)	(-1.25)
Manufacture of textiles, clothing; apparel industry	0.000	0.001
	(0.08)	(1.11)
Manufacture of leather, fur, feather and related products; footwear industry	-0.002**	-0.002**
	(-2.46)	(-2.01)
Processing of timber, manufacture of wood, bamboo, rattan, palm, and straw products	-0.001	-0.001
	(-1.15)	(-0.56)
Manufacture of furniture	-0.002***	-0.002**
	(-2.58)	(-2.25)
Manufacture of paper and paper prod.	-0.002***	-0.001**
	(-2.67)	(-2.03)
Printing and recorded media	-0.001	-0.001
	(-1.45)	(-1.12)
Manufacture of articles for culture, education, art, sports, and entertainment	-0.001	0.000

(continued on next page)

Table 10 (*continued*)

	FE	FGLS
	(-0.52)	(0.18)
Processing of petroleum, coking, processing of nuclear fuel	-0.004***	-0.003**
	(-2.76)	(-2.42)
Manufacture of chemical raw materials and chemical products	-0.000	0.000
	(-0.22)	(0.48)
Pharmaceutical Industry	0.001*	0.001***
	(1.73)	(2.71)
Manufacture of chemical fibers	0.001	0.001
	(0.69)	(1.19)
Manufacture of rubber and plastics	-0.000	0.000
	(-0.64)	(0.17)
Manufacture of non-metallic mineral products	-0.000	0.000
	(-0.00)	(0.87)
Smelting and processing of ferrous metals	-0.003***	-0.002***
	(-3.38)	(-2.90)
Smelting and processing of non-ferrous metals	-0.000	-0.000
	(-0.66)	(-0.09)
Manufacture of metal products	-0.002***	-0.001**
	(-2.81)	(-2.21)
Manufacture of general-purpose machinery	-0.001**	-0.001
	(-2.09)	(-1.24)
Manufacture of special purpose machinery	-0.000	0.001
	(-0.15)	(0.97)
Manufacture of automobiles	-0.001	-0.001
	(-1.07)	(-1.12)
Manufacture of railway, ships, aerospace and other transportation equipment	-0.001	-0.000
	(-1.00)	(-0.45)
Manufacture of electrical machinery and equipment	0.000	0.001
	(0.49)	(1.01)
Manufacture of computers, communication and other electronic equipment	0.002***	0.002***
	(2.68)	(3.20)
Manufacture of measuring instruments	-0.001**	-0.001
	(-1.98)	(-1.29)
Other manufacturing	0.001	0.001
	(0.63)	(1.47)
Comprehensive use of waste resources	-0.001	-0.001
	(-0.81)	(-0.88)
Production and distribution of electric power and heat power	-0.003***	-0.003***
	(-4.93)	(-4.63)
Production and distribution of gas	-0.002*	-0.001
	(-1.71)	(-0.82)
Production and distribution of tap water	-0.003***	-0.003***
	(-3.72)	(-3.36)
Civil engineering	-0.001	-0.000
	(-1.27)	(-0.37)
Renovation	0.002	0.003*
	(1.24)	(1.65)
Decoration	-0.002**	-0.001*
	(-2.29)	(-1.86)
Wholesale trade	-0.000	0.000
	(-0.35)	(0.24)
Retail trade	-0.001	-0.000
	(-1.07)	(-0.73)
Railway transport	-0.004**	-0.003*
	(-2.00)	(-1.84)
Road transport	-0.004***	-0.003***
	(-4.79)	(-4.42)
Water transport	-0.002***	-0.002***
	(-3.29)	(-2.92)
Air transport	-0.002*	-0.002**
	(-1.76)	(-2.06)
Loading/unloading, removal, and other transport services	-0.001	-0.001
	(-0.96)	(-0.48)
Storage	-0.003**	-0.002*
	(-2.55)	(-1.93)
Postal services	0.001	0.001
	(0.83)	(1.15)
Accommodation	-0.000	-0.000
	(-0.18)	(-0.38)

(*continued on next page*)

Table 10 (*continued*)

	FE	FGLS
Catering	-0.001	-0.001
	(-0.41)	(-0.67)
Telecommunications, radio, television, and satellite transmission services	-0.000	-0.001
	(-0.47)	(-0.58)
Internet and related services	0.001**	0.001**
	(1.98)	(2.30)
Software and information technology services	0.001**	0.001***
	(2.33)	(2.70)
Real estate	-0.002**	-0.001**
	(-2.47)	(-2.38)
Leasing	-0.003*	-0.002
	(-1.65)	(-1.34)
Business Service Industry	-0.001*	-0.001
	(-1.81)	(-1.55)
Research and experimental development	-0.001	-0.001
	(-0.94)	(-0.58)
Professional technical service industry	-0.000	0.000
	(-0.69)	(0.21)
Ecological protection and environmental management	-0.002**	-0.001
	(-2.20)	(-1.62)
Management of public facilities	-0.002**	-0.001*
	(-1.97)	(-1.74)
Education	0.002**	0.003***
	(2.15)	(2.69)
Healthcare	-0.001	-0.001
	(-1.37)	(-0.81)
News and publishing	-0.002**	-0.001*
	(-2.23)	(-1.74)
Radio, film, television, and other audio-visual media	-0.001	-0.001
	(-0.99)	(-1.43)
Culture and arts	-0.000	0.000
	(-0.41)	(0.14)
Sports	-0.003**	-0.003*
	(-2.11)	(-1.72)
MKT_t	1.014***	1.015***
	(184.07)	(184.16)
SMB_t	0.566***	0.571***
	(49.99)	(50.35)
HML_t	-0.0678***	-0.0803***
	(-3.11)	(-3.69)
ϕ_0	-0.00270***	-0.00250***
	(-25.22)	(-23.46)
adj. R-sq	0.488	
F	888.0	
N	68083	68083

Notes: This table provides the results from the estimation of the following regression specification.

$R_{i,t} = \varphi_0 + \varphi_1 SENT_{i,t} + \varphi_2 SENT_{i,t} \times Industry_i + \gamma MKT_t + \delta SMB_t + \eta HML_t + \varepsilon_t$

Dependent variable $R_{i,t}$ represents the return of individual stock *i* on day *t* and $SENT_{i,t}$ represents the sentiment index of individual stock *i* on day *t*. $Industry_i$ is the dummy variable based on the Industry Classification Scheme released by China Securities Regulatory Commission in 2020 Q1. Integrated industry is used as the reference group. Companies of the integrated industry have similar revenue and profit in many industries, and no obvious development planning and share holding background. T-stats are in parentheses below the coefficient estimates. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

References

Al-Awadhi, A.M., Al-Saifi, K., Al-Awadhi, A., Alhamadi, S., 2020. Death and contagious infectious diseases: impact of the COVID-19 virus on stock market returns. J. Behav. Exp. Financ., 100326

Baek, S., Mohanty, S.K., Mina, G., 2020. COVID-19 and stock market volatility: an industry level analysis. Financ. Res. Lett., 101748 https://doi.org/10.1016/j.frl.2020.101748.

Baig, A.S., Butt, H.A., Haroon, O., Rizvi, S.A.R., 2020. Deaths, panic, lockdowns and US equity markets: the case of COVID-19 pandemic. Financ. Res. Lett., 101701 https://doi.org/10.1016/j.frl.2020.101701.

Baker, S.R., Bloom, N., Davis, S.J., Kost, K.J., Sammon, M.C., Viratyosin, T., 2020. The Unprecedented Stock Market Impact of COVID-19. National Bureau of Economic Research.

Donadelli, M., Kizys, R., Riedel, M., 2017. Dangerous infectious diseases: bad news for Main Street, good news for Wall Street? J. Financ. Mark. 35, 84–103.

Fallahgoul, H., 2020. Inside the mind of investors during the COVID-19 Pandemic: evidence from the stocktwits data. arXiv Prepr. arXiv 2004, 11686.

Fama, E.F., 1965. Portfolio analysis in a stable Paretian market. Manage. Sci. 11, 404–419.

Ichev, R., Marinč, M., 2018. Stock prices and geographic proximity of information: evidence from the Ebola outbreak. Int. Rev. Financ. Anal. 56, 153–166.

Liu, H., Manzoor, A., Wang, C., Zhang, L., Manzoor, Z., 2020. The COVID-19 outbreak and affected countries stock markets response. Int. J. Environ. Res. Public Health 17, 2800.

Mazur, M., Dang, M., Vega, M., 2020. COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. Financ. Res. Lett., 101690 https://doi.org/10.1016/j.frl.2020.101690.

Nadeem Ashraf, B., 2020. Stock markets' reaction to COVID-19: cases or fatalities? Res. Int. Bus. Financ, 101249.

Shan, L., Gong, S.X., 2012. Investor sentiment and stock returns: wenchuan earthquake. Financ. Res. Lett. 9, 36–47.

Shehzad, K., Liu, X., Kazouz, H., 2020. COVID-19's disasters are perilous than global financial crisis: a rumor or fact? Financ. Res. Lett. 36, 101669 https://doi.org/10.1016/j.frl.2020.101669.

Sun, Yunchuan, Fang, Mengting, Wang, Xinyu, 2018. A novel stock recommendation system using Guba sentiment analysis, (2018), Volume 22 Issue 3, June 2018, Pages 575-587. Personal and Ubiquitous Computing 22 (3), 575–587. https://doi.org/10.1007/s00779-018-1121-x.

Tao, B., 2010. Analysis of the impact of unconventional and unexpected disasters on Chinese stock market volatility – an empirical study based on unit root test of exogenous structural abrupt changes. Knowl. Econ. 42–43.

Y. Sun, M. Fang, X. Wang and S. Diao, GubaLex: Guba-Oriented Sentiment Lexicon for Big Texts in Finance, 2017 13th International Conference on Semantics, Knowledge and Grids (SKG), Beijing, 2017, pp. 25-32.