systems

MDPI

Article

Media News and Social Media Information in the Chinese Peer-to-Peer Lending Market

Jiaqi Kuang [1,*], Xudong Ji [1], Peng Cheng [1] and Vasileios Bill Kallinterakis [2]

[1] International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou 215000, China
[2] Business School, Durham University, Mill Hill Lane, Durham DH1 3LB, UK
* Correspondence: jiaqi.kuang@alumni.xjtlu.edu.cn

Abstract: This paper uses supervised machine learning (sentiment analysis) to analyze the sentiments of social media information in the P2P lending market. After segmentation, filtering, feature word extraction, and model training of the text information captured by Python, the sentiments of media and social media information were calculated to examine the effect of media and social media sentiments on default probability and cost of capital of peer-to-peer (P2P) lending platforms in China (2015–2019). We find that only positive changes in media and social media sentiment have significantly negative effects on the platform's default probability and cost of capital, while negative changes in sentiment do not have any effects. We conclude the existence of an asymmetric effect of media and social media sentiments in the Chinese peer-to-peer lending market.

Keywords: media sentiment; social media sentiment; asymmetry effect; peer-to-peer lending market

check for updates

Citation: Kuang, J.; Ji, X.; Cheng, P.; Kallinterakis, V.B. Media News and Social Media Information in the Chinese Peer-to-Peer Lending Market. *Systems* **2023**, *11*, 133. https://doi.org/10.3390/systems11030133

Academic Editor: Gianfranco Minati

Received: 31 December 2022
Revised: 24 February 2023
Accepted: 27 February 2023
Published: 1 March 2023

1. Introduction

Business presses and media news agencies play an important role as information intermediaries to disseminate information and alleviate information asymmetry in financial markets [1]. Broadly speaking, media can be classified into two separate types: traditional mass media[1] and new media (i.e., digital interactive media or social media[2]), both of which are the important channels for investors to acquire 'private', or say 'inside', information [2,3].

Existing literature has extensively examined the impacts of traditional media information and social media information on financial markets. Many studies state that traditional media information has a significant effect on the firm's performance or valuation, i.e., to increase stock returns and firm values, predict future earnings, and decrease the cost of capital [1,4–6]. However, other research papers find only a weak effect of traditional media coverage [2,7]. Similarly, the relationship between social media information and financial market performance has also been a hot research topic in recent years. Some studies indicate that the effect of social media information on investors' behavior and/or stock market activities is weak [8,9]. However, other studies state that social media information can significantly increase stock returns, predict market volatility, and enhance stock market performance [3,10–13].

Previous literature [3–6,13] states that news and sentiment have a significant effect on firms' performance (trading volume and return) because they contain inside information and market (investor) expectations that help affect firms' performance (trading volume and return). Compared with traditional financial markets (stock and bond markets), innovative and less regulated markets (such as the P2P lending market) present a more serious information asymmetry problem. In addition, as a market whose performance is extremely affected by investors' behavior, news and sentiment contain more inside information and have more power to affect the platforms' performance in the P2P lending market. Therefore, media information becomes more important to reduce information asymmetry. Investors more heavily rely on news from both traditional and social media to make their decisions, which makes the role of information and news in an innovative market (the peer-to-peer lending market) more obvious than in other markets (the market, the bond market).

Even though the body of empirical studies examining the effect of traditional media news and social media posts on the stock and bond markets is relatively mature, in the newly innovative financial markets such as the peer-to-peer (P2P) market, such an effect has not yet been fully studied. However, in an unregulated market, media information becomes more important to reduce information asymmetry. Investors more heavily rely on news from both traditional and social media to make their decisions, which makes the role of information and news in an innovative market (the peer-to-peer lending market) more obvious than in other markets (the stock market, the bond market). We fill a gap and add another piece of empirical evidence to the literature, by examining the effect of media news and social media information on the default probability[3] and cost of capital of China's peer-to-peer online lending platforms. Importantly, we go further to explore the issue of why media news and social media information could have an effect on platforms' performances. According to prior literature [14–18], one possible reason is that sentiment influences investors' behavior. The improved sentiment could enhance investors' confidence and participation; therefore, such reports will help to decrease the default probability and cost of capital of platforms.

In this research, we examine the effect of media and social media sentiments on the default probability and cost of capital of China's peer-to-peer lending platforms. The results show the asymmetry effect between improving and deteriorating sentiment on default probability and cost of capital after using the PSM method[4]. The results prove that only improved media and social media sentiment could help reduce the default probability and cost of capital. We also find a significant and positive effect of a positive change in media/social media sentiment on investor numbers, which is consistent with our argument that the sentiment has affected investors' participation in the P2P market.

Our study contributes to the existing literature in several ways. First, this paper contributes to the emerging literature on the peer-to-peer lending market. It has been widely presented in prior literature that there is a significant information asymmetry between online borrowers and investors in the peer-to-peer lending market [19–24]. Second, our paper contributes to the media news literature [1,2,4–6] and social media information literature [3,10–13,25,26], by extending the media sentiment and social media sentiment effects to the context of the peer-to-peer lending market in China. Importantly, this paper discovers a unique asymmetry effect in the peer-to-peer lending market by demonstrating that only improving sentiments have a significant effect on performance of platforms, but deteriorating sentiments do not. Lastly, this paper goes further to explore why such an effect of media news and social media information on the alternative lending market exists, with the findings revealing that investor behavior plays an important role in these effects.

2. Literature Review and Hypothesis Development

2.1. The Effect of Media News

The media's effect on financial markets, especially stock markets, has been studied extensively [2,4–6,27]. Some papers [1,4–6,28–33] state that there is a significant effect of media news on market performance. Other papers [2,7] hold some different opinions and state that media news may not work for all the firms' performance in the financial market. Even though there are many papers above studying the effect of media coverage on firms' market activities, most of these studies focus on the effect of media news on the stock market and fund market, and there is less research on the effect of media news on the peer-to-peer lending market. Meanwhile, most studies research the effect of media news on firm value, trading volume, and return in the stock market, but few papers study the effect on business bankruptcy and default probability, not to mention in less developed markets such as the peer-to-peer lending market. Therefore, my first hypothesis is stated as follows:

H1: *Media news significantly affects the performance (default probability; cost of capital) of peer-to-peer lending platforms.*

2.2. The Effect of Social Media Information

The economy is a complex system of human interactions [34]. Unlike traditional mass media (i.e., one-way communication in nature), new media, conveying information and stimulating interactions and spreading emotions from one to the other, affect how people feel and drive how they will act. Some papers [3,10–12,35] state there is a significant correlation between social media postings and financial market performance. On the other hand, some papers [8,9] hold different opinions and find there is a weak relationship between social media postings and the stock performance. A few researchers have studied the peer-to-peer lending market as well. Ge et al. [13] state that borrowers' social media account information disclosure (e.g., borrowers' followers) could help reduce borrowers' default probabilities in China. Even though earlier papers studied the effect of social media information, there are few papers focusing on the effect of social media information on default probability at the firms or platform level, especially in the debt market and peer-to-peer lending market. Therefore, my second hypothesis is stated as follows:

H2: *Social media information significantly affects the performance (default probability; cost of capital) of peer-to-peer lending platforms.*

2.3. The Investors' Behavior

Even though there are many papers investigating the media news and social media information effect on financial markets, few papers study the reasons behind the effect. One of the reasons I find that could help explain this effect is that the news sentiment motivates investors' participation, which may be due to the herding effect or the increase in investor recognition. According to prior studies [14–16,18,36], investor sentiment has a significant effect on investors' behavior. In the P2P lending market, the improving sentiment may increase the investors' recognition[5] [37–42] and bring about a herding effect[6] [34,43–46], which will attract more P2P investors or encourage investors to invest more. Meanwhile, the emergence of a large number of active investors is likely to reduce the platform's probability of default and to decrease the cost of capital, while the loss of a large number of investors can increase the probability of default and the cost of capital. Therefore, my third hypothesis is stated as follows:

H3: *Media news and social media information have a significant impact on investors' participation in the peer-to-peer lending market.*

3. Sentiment Analysis

One of the important parts of this paper is the sentiment analysis of media news and social media posts. To examine the media and social media sentiment tendencies, we use the Naive Bayes[7] model in traditional machine learning[8]. We first clear the data by filtering it through the Chinese dictionaries in Snownlp, which is a popular natural language process with generalizing class libraries that was written in Python and used to deal with Chinese text sentiment analysis. Snownlp brings some trained dictionaries that cover most of the Chinese language, and they can be used in different scenes and areas, especially comments and opinions. Many studies [47,48] use it to run the text analysis. Meanwhile, it can also be used to train models by putting specific feature words into the sentiment analysis processing libraries and therefore training them. Then, splitting the words, dropping the low-related words, and keeping the high-related words means the machine needs to clear the common nouns and prepositions. Then, we choose positive feature words and negative feature words though a 1000-person random sample and put these feature words into the model in Snownlp to train. After the training, the model can judge each news item automatically in Python. It will output the probability of the news, which ranges from 0 to 1. If the news probability is not higher than 0.33[9], we judge it as news with a negative sentiment tendency and use −1 to represent it; if the news probability is not higher than 0.66, we judge it as news with a neutral sentiment tendency and use 0 to represent it; and if not, the news should have a positive sentiment tendency and use 1 to represent it. Many people think that the news item's title should have more weight compared with other sentences in the item or post. So, based on our prior study [49], we gave 30% weight to the title and 70% weight to the other content of all news and posts. At last, since we use monthly data to examine the effect of the media news and social media posts, following the method of the previous paper [1], we calculate the monthly media sentiment (*MF*) of each platform by using the aggregated number of positive sentiment news (*N.POS*) minus the aggregated number of negative sentiment news (*N.NEG*) and then dividing by the total amount of news (*N.TOL*) for each platform (*i*) in each month (*t*), the function being as follows:

$$MF_{it} = \frac{N.POS_{it} - N.NEG_{it}}{N.TOL_{it}}$$

4. Methods

To test the effect of media news and social media posts on the default probability at the platform level, the change in media sentiment and social media sentiment are used as testing variables in Model 1. Model 1 is the Logistic Regression, which is one of the linear regressions in which the dependent variable is a binary variable (default contains 2 values: 0 represents survival and 1 represents default). The testing variables are the change in media sentiment ($MCMF_{it} - SMMF_{it\,-\,1}$) and social media sentiment ($SMMF_{it} - SMMF_{it\,-\,1}$). According to the prior studies [50,51], CC^{10} (cost of capital), CR (cumulative repay), ALT (average lending time), NCI (net capital inflow), B (background), and L (location) are included as control variables in Model 1. Model 1 is as follows:

$$DEFAULT_{it} = \alpha_{it} + \beta_1(MCMF_{it} - MCMF_{it-1}) + \beta_2(SMMF_{it} - SMMF_{it-1}) + \sum_{n=3}^{n} \beta_n Controls_{it} + \varepsilon_{it} \quad (1)$$

Model 2 is the Panel Regression, which can solve the problem of missing variables to some extent. In addition, to mitigate the endogeneity problem, we control the time in the random effect model of panel data to make it a two-way fixed effect model. The dependent variable is *CC*, and the testing variables are the change in media sentiment ($MCMF_{it} - SMMF_{it\,-\,1}$) and social media sentiment ($SMMF_{it} - SMMF_{it\,-\,1}$). Based on prior literature [50,52], RF (risk-free rate), CR (cumulative repay), ALT (average lending time), NCI (net capital inflow), B (background), and L (location) are included in our model. To test the effect of media news and social media information on the cost of capital at platform level, the Model 2 is as follows:

$$CC_{it} = \alpha_{it} + \beta_1(MCMF_{it} - MCMF_{it-1}) + \beta_2(SMMF_{it} - SMMF_{it-1}) + \sum_{n=3}^{n} \beta_n Controls_{it} + \varepsilon_{it} \quad (2)$$

As we stated in the literature review [14–16,18,36], one of the reasonable explanations for why media and social media sentiment could affect default probability and a firm's value is that the sentiment could affect investors' participation. To test whether media news and social media information have an effect on the investors behavior, Model 3 is used, which is also a two-way fixed effect model. *IN* is the dependent variable, which is the number of investors on a P2P platform. The change in media sentiment ($MCMF_{it} - MCMF_{it\,-\,1}$) and the change in social media sentiment ($SMMF_{it} - SMMF_{it\,-\,1}$) are used as testing variables.

$$IN_{it} = \alpha_{it} + \beta_1(MCMF_{it} - MCMF_{it-1}) + \beta_2(SMMF_{it} - SMMF_{it-1}) + \sum_{n=3}^{n} \beta_n Controls_{it} + \varepsilon_{it} \quad (3)$$

In order to solve the problems of data bias, variable omission, confounding variables, and sign influence, we use the PSM method to process the above models. Propensity Score Matching (PSM) is a statistical method used to process data. In the observation study, there are many biases and confounding variables. The method of propensity score matching is to reduce the impact of these biases and confounding variables so as to make a more reasonable comparison between the experimental group and the control group. Through data modeling, PSM fits probabilities for each user (multi-dimensional characteristics fit into one-dimensional probability) and searches for the closest sample from the control group and the experimental group for comparison.

5. Results and Discussion

5.1. PSM Results-Default Probability

Previous research has proven the asymmetry effect between goods and bad news on stock volatility [53–55] and return [56,57]. Therefore, we test the different effects of positive and negative sentiment. Table 1 shows the PSM results between the sample observations with positive change in media sentiment and the PSM matched observations without change in media sentiment (listed in (1-1)); the observations with negative change on media sentiment and PSM matched observations without change in media sentiment (listed in (1-2)); and the observations with positive/negative social media sentiment and the PSM matched observations without change in social media sentiment (listed in (1-3) and (1-4)).

Table 1. PSM Results—News Effect on Default.

DEFAULT	(1-1)	(1-2)	(1-3)	(1-4)
PSM-DM	−0.0215 ***	0.0049		
	(0.0080)	(0.0093)		
PSM-DS			−0.0478 **	−0.0337
			(0.0189)	(0.0217)
CC	0.182 ***	0.167 ***	0.127 ***	0.193 ***
	(0.0141)	(0.0175)	(0.0266)	(0.0323)
CR	0.0367 ***	0.0465 ***	0.0179 ***	0.0037
	(0.0044)	(0.0075)	(0.0062)	(0.0095)
ALT	−0.148 ***	−0.138 ***	−0.0927 ***	−0.0990 ***
	(0.0123)	(0.0181)	(0.0238)	(0.0261)
NCI	0.0003	−0.0013	−0.0028	−0.0035 *
	(0.0012)	(0.0017)	(0.0020)	(0.0021)
T	YES	YES	YES	YES
B	YES	YES	YES	YES
L	YES	YES	YES	YES
Constant	−4.8052 ***	−3.3804 ***	−3.1227 ***	−2.1745 ***
	(0.6750)	(0.3824)	(0.4986)	(0.6130)
Observations	4331	2433	1251	1061
No. Platforms	970	541	252	231
Prob > chi^2	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.0436	0.0363	0.0444	0.0617
Observations	4275	2328	3576	1653
No. Platforms	970	553	969	464

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is DEFAULT. Model (1-1) compares the positive DMCMF (set as 1 in PSM-DM) and 0 DMCMF (set as 0 in PSM-DM); and model (1-2) compares the negative DMCMF (set as 1 in PSM-DM) and 0 DMCMF (set as 0 in PSM-DM); Model (1-3) compares the positive DSMMF (set as 1 in PSM-DS) and 0 DSMMF (set as 0 in PSM-DS), and model (1-4) compares the negative DSMMF (set as 1 in PSM-DS) and 0 DSMMF (set as 0 in PSM-DS). The results show that only positive DMCMF and positive DSMMF have significant effects on decreasing platform defaults.

In models (1-1) and (1-3), the significant coefficients of PSM-DM (positive DMCMF) and PSM-DS (positive DSMMF) indicate that increasing media sentiment (MCMF) and social media sentiment (SMMF) help reduce the default probability. In models (1-2) and (1-4), the insignificant PSM-DM (negative DMCMF) and PSM-DS (negative DSMMF) show no effect of the negative change in media sentiment (MCMF) and social media sentiment (SMMF) on default probability.

These results demonstrate in H1 and H2 that media news and social media information significantly affect the default probability of peer-to-peer lending platforms and also indicate that only positive changes in media/social media sentiment have a significant impact on decreasing default. The possible explanation is that the positive changes in media/social media sentiment have been interpreted as a good indicator by investors, who are subsequently willing to invest more capital into platforms. On the other hand, the negative changes in media/social media sentiment have been interpreted as a bad signal by investors; however, investors cannot leave immediately because all the bids and products in the peer-to-peer lending market cannot be sold or withdrawn in a short period of time once the investment is made. Another possible explanation is that most investors in innovative financial markets (such as the peer-to-peer lending market) are risk lovers rather than risk averters; therefore, they are more sensitive to improving sentiments and less sensitive to deteriorating sentiments. Our results demonstrate the existence of an asymmetry effect between positive sentiment and negative sentiment in the peer-to-peer lending market, which may have practical implications for regulators in terms of how such market shall be regulated.

5.2. PSM Results—Cost of Capital

Table 2 compared the results of samples with a positive/negative change of media sentiment and a PSM matched sample with zero change of media sentiment (listed in (2-1)/(2-2)); and the results of a sample with positive/negative change of social media sentiment to the results of the PSM matched sample with zero change in social media sentiment (listed in (2-3)/(2-4)).

Table 2. PSM Results—News Effect on Cost of Capital.

CC	(2-1)	(2-2)	(2-3)	(2-4)
PSM-RM	−0.0335 **	0.0174		
	(0.0136)	(0.0127)		
PSM-RS			−0.0482 ***	−0.0179
			(0.0145)	(0.0231)
RF	9.211 ***	1.746 ***	9.416 ***	2.065 ***
	(0.641)	(0.236)	(0.663)	(0.453)
CR	−0.0589 ***	−0.0617 ***	−0.0615 ***	−0.0815 ***
	(0.0053)	(0.0055)	(0.0055)	(0.0094)
ALT	0.570 ***	0.899 ***	0.568 ***	0.640 ***
	(0.0104)	(0.0071)	(0.0108)	(0.0139)
NCI	−0.0026 **	−0.0019 **	−0.0023 *	−0.0029 *
	(0.0011)	(0.0009)	(0.0012)	(0.0017)
T	YES	YES	YES	YES
B	YES	YES	YES	YES
L	YES	YES	YES	YES
Constant	−8.320 ***	−8.311 ***	−8.489 ***	−8.467 ***
	(0.749)	(0.747)	(0.770)	(0.766)
Prob > chi^2	0.0000	0.0000	0.0000	0.0000
R-squared	0.4315	0.6688	0.4401	0.4752
Observations	4275	2328	3576	1653
No. Platforms	970	553	969	464

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is CC. Model (2-1) compares the positive DMCMF (set as 1 in PSM-RM) and 0 DMCMF (set as 0 in PSM-RM); and model (2-2) compares the negative DMCMF (set as 1 in PSM-RM) and 0 DMCMF (set as 0 in PSM-RM); Model (2-3) compares the positive DSMMF (set as 1 in PSM-RS) and 0 DSMMF (set as 0 in PSM-RS); and model (2-4) compares the negative DSMMF (set as 1 in PSM-RS) and 0 DSMMF (set as 0 in PSM-RS). The results show that only positive DMCMF and positive DSMMF have a significant effect on decreasing the cost of capital for platforms.

In models (2-1) and (2-3), the significant coefficients of PSM-RM (positive DMCMF) and PSM-RS (positive DSMMF) also indicate that increasing media sentiment (MCMF) and social media sentiment (SMMF) have a significant impact on reducing the cost of capital. In models (2-2) and (2-4), the insignificant PSM-RM (negative DMCMF) and PSM-RS (negative DSMMF) show no effect of the negative change in media sentiment and social media sentiment on the cost of capital. The results listed there demonstrate for hypotheses H1 and H2 that media news and social media information have a significant effect on the cost of capital of peer-to-peer lending platforms. These results also indicate the asymmetry effect between positive and negative changes in sentiment in the peer-to-peer lending market.

5.3. PSM Results—Investors' Behavior

In order to further test our above conjecture about the reasons for the asymmetrical effect of media and social media sentiment, we examine the effect of sentiments on investors' participation in the peer-to-peer lending market. Table 3 compared the results between samples with a positive/negative change in media sentiment and a PSM matched sample with zero change in media sentiment (listed in (3-1)/(3-2)); and the results between samples with positive/negative change of social media sentiment and a PSM matched sample with zero change in social media sentiment (listed in (3-3)/(3-4)). In models (3-1) and (3-3), the significant coefficients of PSM-IM (positive DMCMF) and PSM-IS (positive DSMMF) indicate that increasing media sentiment (MCMF) and social media sentiment (SMMF) help raise investor numbers. In models (3-2) and (3-4), the insignificant PSM-IM (negative DMCMF) and PSM-IS (negative DSMMF) show that negative changes in MCMF and SMMF have no effect on investor numbers. The results prove H3—that media news and social media information have a significant impact on investors' participation in the peer-to-peer lending market. These results are also consistent with our supposition that only increasing (positive change in) sentiments (both media and social media) could attract more investors and boost investor confidence, thereby helping reduce platform default probability and cost of capital.

Table 3. PSM Results—Investors' Behavior.

IN	(3-1)	(3-2)	(3-3)	(3-4)
PSM-IM	0.192 ***	0.0783		
	(0.0426)	(0.0611)		
PSM-IS			0.216 ***	0.0686
			(0.0493)	(0.0736)
CC	1.488 ***	1.436 ***	1.519 ***	1.450 ***
	(0.0480)	(0.0694)	(0.0539)	(0.0834)
CR	0.691 ***	0.549 ***	0.709 ***	0.462 ***
	(0.0165)	(0.0214)	(0.0176)	(0.0236)
ALT	−0.00489	0.240 ***	−0.0457	0.436 ***
	(0.0438)	(0.0612)	(0.0481)	(0.0713)
NCI	0.0331 ***	0.0268 ***	0.0309 ***	0.0368 ***
	(0.0037)	(0.0051)	(0.0042)	(0.0059)
T	YES	YES	YES	YES
B	YES	YES	YES	YES
L	YES	YES	YES	YES
Constant	−5.398 ***	−2.957 ***	−5.548 ***	−2.942 **
	(0.614)	(1.045)	(0.635)	(1.189)
Prob > chi^2	0.0000	0.0000	0.0000	0.0000
R-squared	0.7460	0.7691	0.7320	0.7648
Observations	4275	2328	3576	1653
No. Platforms	970	553	969	464

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable is IN. Model (3-1) compares the positive DMCMF (set as 1 in PSM-IM) and 0 DMCMF (set as 0 in PSM-IM); and model (3-2) compares the negative DMCMF (set as 1 in PSM-IM) and 0 DMCMF (set as 0 in PSM-IM); Model (3-3) compares the positive DSMMF (set as 1 in PSM-IS) and 0 DSMMF (set as 0 in PSM-IS), and model (3-4) compares the negative DSMMF (set as 1 in PSM-IS) and 0 DSMMF (set as 0 in PSM-IS). The results show that only positive DMCMF and positive DSMMF have a significant effect on increasing the number of platforms with investors.

6. Conclusions

This study investigates the effects of media and social media sentiment on default probability and cost of capital in the Chinese peer-to-peer lending market. Using the unique media news and social media posts dataset that was collected by Python and analyzed by Snownlp, a sentiment analysis instrument, we find that only the positive changes in media and social media sentiment could reduce the default probability and cost of capital, while the negative changes in sentiment had no effect. Furthermore, the media and social media sentiment could affect investors' participation and behavior in the P2P lending market because positive changes in sentiment can attract more investors.

This study has some implications for participants in the peer-to-peer lending market. For policymakers, the government should focus on the effect of media news and social media posts in the innovation market by monitoring the media sentiment and social media sentiment in the market. This will help improve the governments regulation of the newly established financial market. Investors should pay greater attention to platforms that improve media and social media sentiment because they will help reduce default probability and cost of capital (which is also return for investors).

Author Contributions: Conceptualization, J.K. and X.J.; methodology, J.K.; software, J.K.; validation, J.K., X.J., P.C., and V.B.K.; formal analysis, J.K. and X.J.; resources, J.K., X.J., and P.C.; supervision, X.J., P.C., and V.B.K.; funding acquisition, X.J. and P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Xi'an Jiaotong-Liverpool Univesity [Ref: RDF-15-02-04].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data is grabbed by the author from the Chinese Internet with data mining.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Media news is the news published by traditional media or mass media, such as the Wall Street Journal or China Daily.
2. Social media information is the news or information published by social media or new media, such as Twitter, Facebook, and Weibo.
3. Default probability: this is a dummy variable with a value of 1 if the platform defaults and 0 otherwise.
4. Propensity Score Matching (PSM) is a statistical method used to process data. Through data modeling, PSM fits probabilities for each user (multi-dimensional characteristics fit into one-dimensional probability) and searches for the closest sample from the control group and the experimental group for comparison.
5. Investors will only invest in securities they know about. If a company is known by more investors, it will reduce information asymmetry [36].
6. Herding behavior occurs when a group of investors intentionally follows the actions or reactions of other investors whom they consider to be better informed, instead of following their own beliefs and using their own information when they make the decisions [49].
7. The naive Bayes method is a classification method based on the Bayes theorem and independent hypotheses of feature conditions. The naive Bayesian algorithm is widely used in text recognition, text classification, and image recognition. It can classify an unknown text or image according to its existing classification rules and finally achieve the purpose of classification.
8. We also use the BP (back propagation) model, which is a widely used neural network, to run the robustness check. All the results are similar. The results are available if asked.
9. The 0.33 and 0.66 are set in Python and Snownlp, based on the previous paper [29], the values 0.33 and 0.66 should be used when we code the sentiment analysis.
10. The CC (cost of capital) here also represents the return from an investors' perspective.

References

1. Cahan, F.S.; Chen, C.; Chen, L.; Nguyen, H.N. Corporate Social Responsibility and Media Coverage. *J. Bank. Financ.* **2015**, *59*, 409–422. [CrossRef]
2. Fang, L.; Peress, J. Media Coverage and the Cross-section of Stock Returns. *J. Financ.* **2009**, *64*, 2023–2052. [CrossRef]
3. Da, Z.; Engelberg, J.; Gao, P. In Search of Attention. *J. Financ.* **2011**, *66*, 1461–1499. [CrossRef]
4. Barber, B.M.; Loeffler, D. The "Dartboard" Column: Second-hand Information and Price Pressure. *J. Financ. Quant. Anal.* **1993**, *28*, 273–284. [CrossRef]
5. Albert, R.L., Jr.; Smaby, T.R. Market Response to Analyst Recommendations in the "Dartboard" Column: The Information and Price-pressure Effects. *Rev. Financ. Econ.* **1996**, *5*, 59–74. [CrossRef]
6. Tetlock, P.C.; Tsechshansky, M.S.; Macskassy, S. More Than Words: Quantifying Language to Measure Firm' Fundamentals. *J. Financ.* **2008**, *63*, 1437–1467. [CrossRef]
7. Wang, J.; Ye, K. Media Coverage and Firm Valuation: Evidence from China. *J. Bus. Ethics* **2015**, *127*, 510–511. [CrossRef]
8. Tumarkin, R.; Whitelaw, R.F. News or Noise? Internet Postings and Stock Prices. *Financ. Anal. J.* **2001**, *57*, 41. [CrossRef]
9. Dewally, M. Internet Investment Advice: Investing with a Rock of Salt. *Financ. Anal. J.* **2003**, *59*, 65–77. [CrossRef]
10. Antweiler, W.; Frank, M.Z. Is All that Talk Just Noise? The Information Content of Internet Stock Message Boards. *J. Financ.* **2004**, *59*, 1259–1294. [CrossRef]
11. Das, S.R.; Chen, M.Y. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Manag. Sci.* **2007**, *53*, 1375–1388. [CrossRef]
12. Bollen, J.; Mao, H.; Zeng, X. Twitter Mood Predicts the Stock Market. *J. Comput. Sci.* **2011**, *2*, 1–8. [CrossRef]
13. Ge, R.; Feng, J.; Gu, B.; Zhang, P. Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending. *J. Manag. Inf. Syst.* **2017**, *34*, 401–424. [CrossRef]
14. Luo, J.S.; Li, C.A. Futures Market Sentiment and Institutional Investor Behavior in the Spot Market: The Emerging Market in Taiwan. *Emerg. Mark. Financ. Trade* **2008**, *44*, 70–86. [CrossRef]
15. Gao, B.; Yang, C. Investor Trading Behavior and Sentiment in Futures Markets. *Emerg. Mark. Financ. Trade* **2018**, *54*, 707–720. [CrossRef]
16. Kim, K.; Ryu, D. Does Sentiment Determine Investor Trading Behavior? *Appl. Econ. Lett.* **2021**, *28*, 811–816. [CrossRef]
17. Jiang, Y.; Ho, Y.C.; Yan, X.; Tan, Y. Investor Platform Choice: Herding, Platform Attributes, and Regulations. *J. Manag. Inf. Syst.* **2018**, *35*, 86–116. [CrossRef]
18. Chiang, M.T.; Lin, M.C. Market Sentiment and Herding in Analysts' Stock Recommendations. *North American J. Econ. Financ.* **2019**, *48*, 48–64. [CrossRef]
19. Berkovich, E. Search and Herding Effects in Peer-to-peer Lending: Evidence from Prosper.com. *Ann. Financ.* **2011**, *7*, 389–405. [CrossRef]
20. Herzenstein, M.; Sonenshein, S.; Dholakia, M.U. Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-peer Lending Decisions. *J. Mark. Res.* **2011**, *48*, S138–S149. [CrossRef]
21. Michels, J. Do Unverifiable Disclosure Matter? Evidence from Peer-to-peer Lending. *Account. Rev.* **2012**, *87*, 1385–1413. [CrossRef]
22. Duarte, J.; Siegel, S.; Young, L. Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *Rev. Financ. Stud.* **2012**, *25*, 2455–2484. [CrossRef]
23. Chen, J.; Jiang, J.; Liu, Y. Financial Literacy and Gender Difference in Loan Performance. *J. Empir. Financ.* **2018**, *48*, 307–320. [CrossRef]
24. Lin, M.; Prabhala, P.N.; Viswanathan, S. Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-peer Lending. *Manag. Sci.* **2018**, *59*, 17–35. [CrossRef]
25. Mai, F.; Shan, Z.; Bai, Q.; Wang, X.; Chiang, R.H. How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis. *J. Manag. Inf. Syst.* **2018**, *35*, 19–52. [CrossRef]
26. Xu, J.J.; Chau, M. Cheap talk? The Impact of Lender-Borrower Communication on Peer-to-peer Lending Outcomes. *J. Manag. Inf. Syst.* **2018**, *35*, 53–85. [CrossRef]
27. Joseph, N.; Kar, A.K.; Ilavarasan, P.V.; Ganesh, S. Review of Discussions on Internet of Things (IoT): Insights from Twitter Analytics. *J. Glob Inf. Manag.* **2017**, *25*, 38–51. [CrossRef]
28. Ferreira, E.J.; Smith, S.D. Stock Price Reactions to Recommendations in the Wall Street Journal "Small Stock Focus" Column. *Q. Rev. Econ. Financ.* **1999**, *39*, 379–389. [CrossRef]
29. Kerl, A.G.; Walter, A. Market Responses to Buy Recommendations Issued by Personal Finance Magazines: Effects of Information, Price-pressure, and Company Characteristics. *Rev. Financ.* **2007**, *11*, 117–141. [CrossRef]
30. Engelberg, J.E.; Parsons, C.A. The Causal Impact of Media in Financial Markets. *J. Financ.* **2011**, *66*, 67–97. [CrossRef]
31. Dougal, C.; Engelberg, J.; García, D.; Parsons, C.A. Journalists and the Stock Market. *Rev. Financ. Stud.* **2012**, *25*, 639–679. [CrossRef]
32. Peress, J. The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes. *J. Financ.* **2014**, *69*, 2007–2043. [CrossRef]
33. Ozik, G.; Sadka, R. Media Coverage and Hedge Fund Returns. *Financ. Anal. J.* **2013**, *69*, 57–75. [CrossRef]
34. Nofsinger, J.R. Social Mood and Financial Economics. *J. Behav. Financ.* **2005**, *6*, 144–160. [CrossRef]

35. Sul, H.K.; Dennis, A.R.; Yuan, L. Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns. *J. Decis. Sci. Inst.* **2016**, *48*, 454–488. [CrossRef]
36. Hudson, Y.; Yan, M.; Zhang, D. Herd Behavior & Investor Sentiment: Evidence from UK Mutual Funds. *Int. Rev. Financ. Anal.* **2020**, *71*, 101494. [CrossRef]
37. Agmon, T.; Donald, R.L. Investor recognition of corporate international diversification. *J. Financ.* **1977**, *32*, 1049–1055. [CrossRef]
38. Merton, R.C. A Simple Model of Capital Market Equilibrium with Incomplete Information. *J. Financ.* **1987**, *42*, 483–510. Available online: https://dspace.mit.edu/bitstream/handle/1721.1/2166/SWP-1869-18148074.pdf (accessed on 26 February 2023). [CrossRef]
39. Foerster, S.R.; Karolyi, G.A. The Effects of Market Segmentation and Investor Recognition on Asset Prices: Evidence from Foreign Stocks Listing in the United States. *J. Financ.* **2002**, *54*, 981–1013. [CrossRef]
40. Bodnaruk, A.; Ostberg, P. Does Investor Recognition Predict Returns? *J. Financ. Econ.* **2009**, *91*, 208–226. [CrossRef]
41. Green, T.C.; Jame, R. Company Name Fluency, Investor Recognition, and Firm Value. *J. Financ. Econ* **2013**, *109*, 813–834. [CrossRef]
42. Jacobs, H.; Hillert, A.; Bias, A. Investor Recognition, and Trading Behavior. *Rev. Financ.* **2016**, *20*, 693–723. [CrossRef]
43. Sias, R.W. Institutional Herding. *Rev. Financ. Stud.* **2004**, *17*, 165–206. [CrossRef]
44. Blasco, N.; Ferreruela, S. Testing Intentional Herding in Familiar Stocks: An Experiment in an International Context. *J. Behav. Financ.* **2008**, *9*, 72–84. [CrossRef]
45. Chen, T. Investor Attention and Global Stock Returns. *J. Behav. Financ.* **2017**, *18*, 358–382. [CrossRef]
46. Chong, T.; Liu, X.; Zhu, C. What Explains Herd Behavior in the Chinese Stock Market? *J. Behav. Financ.* **2017**, *18*, 448–456. [CrossRef]
47. Chen, C.; Chen, J.; Shi, C. Research on Credit Evaluation Model of Online Store Based on SnowNLP. In Proceedings of the 3rd International Conference on Advances in Energy and Environment Research (ICAEER), Guilin, China, 10–12 August 2018; Volume 53. [CrossRef]
48. Zhao, Z.S.; Jin, X.S.; Wen, T.T.; Liang, J.H. News Comment Data Analysis Based on Python-Snownlp. *Public Commun. Sci. Technol.* **2018**, *10*, 104–105. Available online: https://www.cnki.com.cn/Article/CJFDTotal-KJCB201818057.htm (accessed on 26 February 2023).
49. Piotroski, D.J.; Wong, J.T.; Zhang, T. Political Bias in Corporate News: The Role of Conglomeration Reform in China. *J. Law Econ.* **2017**, *60*, 173–207. [CrossRef]
50. Carlos, S.C.; Begoña, G.N.; Luz, L.P. Determinants of Default in P2P Lending. *PLoS ONE* **2015**, *10*, e0139427. [CrossRef]
51. Wang, Q.; Si, Z.; Chen, X. Information Disclosure and the Default Risk of Online Peer-to-peer Lending Platform. *Financ. Res. Lett.* **2020**, *38*, 101509. [CrossRef]
52. Fledstein, M.; Eckstein, O. The Fundamental Determinants of the Interest Rate. *Rev. Econ. Stat.* **1970**, *52*, 363–375. [CrossRef]
53. Engle, R.F.; Ng, V.K. Measuring and Testing the Impact of News on Volatility. *J. Financ.* **1993**, *48*, 1749–1778. [CrossRef]
54. Braun, P.A.; Nelson, D.B.; Sunier, A.M. Good News, Bad News, Volatility, and Betas. *J. Financ.* **1995**, *50*, 1575–1603. [CrossRef]
55. Malik, F. Estimating the Impact of Good News on Stock Market Volatility. *Appl. Financ. Econ.* **2011**, *21*, 545–554. [CrossRef]
56. Bae, K.H.; Karolyi, G.A. Good News, Bad News and International Spillovers of Stock Return Volatility between Japan and the US. *Pac. -Basin Financ. J.* **1994**, *2*, 405–438. [CrossRef]
57. Depken, G.A. Good news, Bad News and GARCH Effects in Stock Return Data. *J. Appl. Econ.* **2011**, *4*, 313–327. [CrossRef]