PACIS 2019 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

6-15-2019

Not only Online Review but also its Helpfulness is Manipulated: Evidence from Peer to Peer Lending Forum

Liting Li
Southwestern University of Finance and Economics, 664922567@qq.com

Haichao Zheng
Southwestern University of Finance and Economics, haichao@swufe.edu.cn

Dongyu Chen
Soochow University, 33671446@qq.com

Bin Zhu
Oregon State University, bin.zhu@bus.oregonstate.edu

Not only Online Review but also its Helpfulness is Manipulated: Evidence from Peer to Peer Lending Forum

Completed Research Paper

Liting Li
Southwestern University of Finance and Economics
No. 555 Liutai Ave., Wenjiang, Chengdu, China
664922567@qq.com

Haichao Zheng
Southwestern University of Finance and Economics
No. 555 Liutai Ave., Wenjiang, Chengdu, China
haichao@swufe.edu.cn

Dongyu Chen
Soochow University
No.50 Donghuan road, Suzhou, China
chendongyu@suda.edu.cn

Bin Zhu
Oregon State University
Austin Hall 422, 2751 SW Jefferson Way, Corvallis, OR, United States
bin.zhu@bus.oregonstate.edu

Abstract

Online reviews have become proposed as useful information for consumers to make decision. Meanwhile, review manipulation will weaken the credibility of online reviews. Except manipulating the review text and rating, we propose that review helpfulness, an important signal for consumer to filter the reviews, could also be manipulated. This study thus explores the existence of review helpfulness manipulation and the relationship between firm quality and review manipulation. Based on a dataset from a review forum in www.wdzj.com which is the leading and largest portal of peer to peer lending industry in China, we get the following interesting results. First, due to the manipulation of review helpfulness, a manipulated positive review is more likely to receive higher helpfulness, while a manipulated negative is more likely to get lower helpfulness. Second, a manipulated review tends to be lower quality in terms of readability and word count, which are found as positive predictors for review helpfulness. Third, high quality firms tend to manipulate more positive reviews, and at the same time high quality firms will receive more negative manipulated reviews. This study extends current understanding about online review manipulation, thereby providing theoretical and practice implications.

Keywords: review manipulation, review helpfulness, review quality, firm quality

Introduction

Online review forums like Yelp (www.yelp.com) and Dianping (www.dianping.com) offer new channels for consumers to communicate and express their opinions on a product (Cheung and Thadani, 2012). Expressing electronic WOM could satisfy consumers' desire for social interaction and economic incentives, express their concern for other consumers, and have the potential to enhance their own self-worth. Some studies in marketing and electronic commerce literature indicate that online reviews of product or service could acts as a signal of product quality which will help consumers make better choices (Wang and Yu, 2015). However, other scholars also question the informativeness of online reviews, as the online reviews might be manipulated by the firms (Hu et al., 2011a).

Following Hu et al. (2011a), we define the review manipulation (or fraud) as occurring when online firms, publishers, or even their competitors write “consumers” reviews by posing as real customers. Reviews manipulation has been identified as a popular practice in music industry in which professional marketers are hired to post positive comments about new albums in many online forums (Mayzlin 2006). Such practice also exists in famous electronic commerce giant like Amazon. For instance, Harmon (2004) shows that a proportion of book reviews were written by the books’ own publishers, authors, and competitors.

Online review manipulation also gained much attention from scholars and online forum operators (e.g., Dellarocas 2006; Hu et al., 2012). Specifically, diverse manipulation detection methods have been developed. For example, Wald–Wolfowitz (Runs) test was proposed by Hu et al. (2012) to test whether or not the reviews for a product follows a random manner, which implies that a non-random pattern indicates the existence of manipulation. Some online forums operators also try to develop technologies or governance mechanisms to increase the cost of review manipulation, thereby reducing firms conduct review manipulation (Dellarocas 2006).

Most of the studies on review manipulation focus on the review rating, readability, and sentiments (e.g., Hu et al., 2012). Current research assumes that helpfulness of reviews is an indicator for the existence of review manipulation. For example, Hu et al. (2011b) suggest that books whose reviews on an average are rated as highly helpful can serve as a non-manipulation indicator. In this study, we make a competing proposition that if the helpfulness is manipulated it will not works as a signal for non-manipulation.

The possible manipulation of the review helpfulness is worthy of investigation for several reasons. First, most of the forums offer a button asking the users “was this review helpful for you?”, which could help subsequent consumers filter informative ones (Korfiatis et al., 2012). It is easier for a manipulator to vote for his/her review by clicking that button. Second, the manipulation of review helpfulness make it is not a valid measurement of the true helpfulness of a review. Thus, researchers should be cautious to use the raw data of review helpfulness like the ratio of the helpful votes over the total number of votes as a variable in the research model (e.g., Korfiatis et al., 2012).

The helpful or non-helpful votes of online reviews may come from two sources – true consumers and the manipulators who pose that review. As a result, a fraud review gets helpfulness votes in two ways. First, the potential true users may give positive or negative votes for the review as in most cases they cannot distinguish the source of review (Dellarocas 2006). So, in order to identify the existence of helpfulness manipulation, we need to distinguish the effects of helpfulness manipulation and the quality of review which affecting true users’ votes. Therefore, our first research question is as follows:

R1: Does a manipulated review affect its helpfulness by means of the mediating role of text quality.

Second, the manipulator directly or indirectly gives positive votes for the review posted by his/her self. For example, the manipulator could register in the online forum as another users and directly provide helpful votes. Except directly manipulating helpfulness, the manipulator could offer money to hire other marketers to give positive votes. To study whether or not the helpfulness manipulation exists, we empirically study the following research questions:

R2: Does firm directly manipulate the helpfulness of online reviews?

Finally, we will further study what types of firms are more likely to have manipulated online reviews and helpfulness. Specifically, “have” here means that manipulated reviews may be the positive comments come from firm itself or the negative comments given by its competitors. Extant research shows that, for the market like music and books where prices are essentially exogenously fixed, low quality firm manipulate its ratings of reviews more than the high firm (Dellarocas 2006; Hu et al., 2011a). Dellarocas (2006) also proposes that in a duopoly setting where prices are endogenous, if the marginal revenues gains are increasing function of consumers’ beliefs about the firm’s quality, the high quality firm will manipulate more than the low quality firm. However, we find that few empirical researches have shown that the manipulation intensity of high quality firm is higher. The peer to peer (P2P) lending industry may offer a setting that the service price could be endogenous set, which could enable us to empirically test whether or not the high quality financing service provider will manipulate more. In addition, competitors may post offensive negative review, i.e., manipulated reviews, to defame the high quality firm. To uncover the manipulation behavior of different types of firm, we will empirically study the following two questions:

R3: Are high quality or low quality firms are more likely to manipulate positive reviews?

R4: Are high quality firms are more likely to receive manipulated negative reviews?

This paper proceeds as follows. The next section develops hypotheses. Then the methodology part presents the research framework, data and measurement. The results part shows the answers to the research questions. Finally, we conclude the paper with discussion, limitation and furfure research directions.

Conceptual Model and Hypotheses Development

The conceptual model is developed as figure 1. Research unit in this study is individual review. The variable review helpfulness refers the ratio of number of helpful vote a review receive divided by the total number of votes. The model in figure 1 shows that whether or not a review is manipulated have effects on helpfulness through direct and indirect mediating routes. Additionally, firm quality is proposed as an antecedent of manipulation.

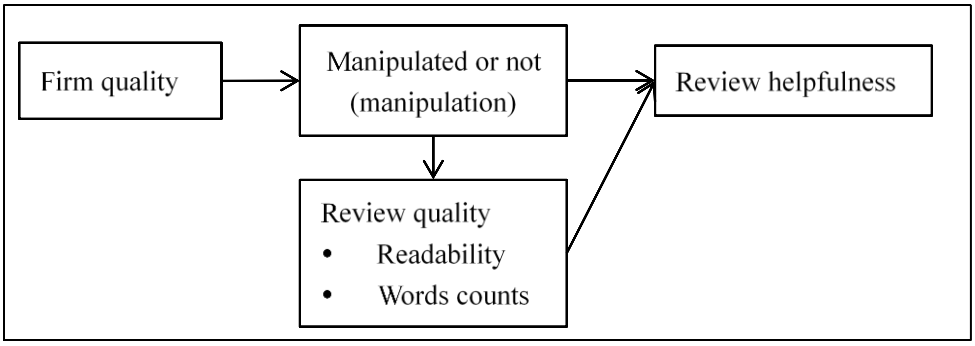

Figure 1 Conceptual Model

Mediating Role of Review Quality

The quality of online review is operationalized as text readability and word counts (Korfiatis et al., 2012). Readability refers to the cognitive effort required for a person to comprehend a piece of text (Zakaluk and Samuels, 1988). Words count, as the term implies, is the total number of words in a review. To study the mediating roles of review quality, we need to test whether or not manipulation has significant effect on review quality, and at the same time review quality affects review helpfulness. According to prior research, a piece of review written in a easy to follow manner will be recognized as more useful review by true consumers, who may give positive votes (Korfiatis et al., 2012). In addition, review depth (word count) has been proved having a positive effect on the helpfulness of the review (Mudambi and Schuff, 2010). So, we make the following hypothesis:

H1: Review quality (readability and words count) will improve its helpfulness.

Prior studies suggest that a fraud review is less readable than a true review (Ong, 2013). For example, Moffitt and Burns (2009) find that fraudulent financial reports are more likely to contain more complex words than the normal ones. One possible reason is that a manipulator who writes a fraud review will try to confuse others by using complex sentence. As for total words count, the manipulator might be unwilling to make efforts write a large piece of text. Thus, we propose the following hypothesis:

H2: A manipulated review tends to be of lower text quality.

Manipulation of Review Helpfulness

To filter low quality online review, most of the forums motive users to vote for the review's helpfulness. We agree with Hu et al. (2011b) that if review helpfulness is not manipulated, it can serve as an indicator of non-manipulation of review. However, we must be caution that, to enable a manipulated review become "useful", the manipulator will be prone to give positive vote directly or hire others to vote for their reviews, thereby increasing their fraud reviews' influence. Additionally, compared with high quality fake review, manipulating review helpfulness is easier and cost less. Therefore, we make the following hypothesis:

H3: A manipulated review is more likely to receive higher helpfulness due to the manipulation of review helpfulness.

Firm Quality and Reviews Manipulation

In the market like books, low quality firm manipulates its ratings of reviews more than the high quality firm (Hu et al., 2011a). The research context in our study is P2P lending which is a type of lending that enables people to make direct borrow and lend via an online platform (Lin et al., 2013). Unlike the traditional market like music and books where prices are essentially exogenously fixed, the service fees are endogenous in P2P lending platforms. In addition, compared with the hotel industry (Zhuang et al., 2018), consumers can hardly get firm's quality information offline. Thus, manipulation of online review has greater effect on consumers' perception of P2P lending platform. In this market, the high quality firms tend to manipulate more, which could accentuates the gap between the ratings of the low quality competitors and itself. So, higher intensity of manipulation from high quality firms could help investors distinguish the high P2P lending platform, thereby increasing high quality firm's revenue (Dellarocas 2006). So we expect that high quality firms tend to manipulate more positive reviews. However, the competition among P2P lending platforms also brings more negative fraud reviews from competitors. Hence, we make the following two hypotheses:

H4: High quality firms tend to manipulate more positive reviews.

H5: High quality firms will receive more negative manipulated review.

Methodology

Data

We collect our data from WDZJ (www.wdzj.com), which is the leading and largest portal of online lending industry in China. It provides all-round and authoritative data of P2P lending platforms in China. All P2P lending platforms are ranked by the WDZJ based on its transactions volume. Except, the lending and loan information, consumers' discussion forum can also give reviews to each platform.

In our study, data on every platform available was gathered automatically with a self-developed web crawler to retrieve forum data. Our forum data covers about 5 years from October 2014 to August 2018, including 202,667 reviews from users among 6276 platforms. Firstly, in our dataset we only include the platforms that receive more than 20 reviews. In addition, the replies of each review are excluded in our data analysis. After removing those platforms and replies, 140,007 review and 738 platforms were left. In addition, some platforms do not have any performance information reported in WDZJ. As the performance data is used as proxy for firms' quality, so finally there are 88,973 reviews and 293 platforms were included in the dataset.

Specially, we will introduce what the review information look like in the review forum of WDZJ. Besides the user's information, a review in WDZJ forum consists of three parts: numerical expressions (helpful votes, helpless votes, numerical ratings which is between 1 to 5 of four dimensions of service quality), an overall evaluation indicating whether or not the reviewer like to recommend the P2P lending platform to other users, and the textual comments about the platform.

Figure 2 shows a review posted on WDZJ. The sum of the four ratings of speed, guard, website experience, and service is used as the total rating of the platform.

Figure 2 An example of review posted on WDZJ

Review Manipulation Detection

The research unit in our study is the individual review. Thus, firstly we must detect whether or not one review is manipulated. As Wald–Wolfowitz (Runs) test or time series method works for detection non-random pattern of all reviews for one product or firm (Hu et al., 2012), we do not adopt that approach. Instead, we try to develop multiply rules to label the review as normal or abnormal. Abnormal review here can be considered as manipulated ones.

We designed three categories of rules to detect abnormal reviews that may be manipulated. As long as a review satisfies at least one rule, we label it as a suspected manipulated one. The rules and number of reviews identified based on the rules are listed in Table 1. Finally, we identified 12,201 suspected manipulated reviews out of 140,007 total reviews.

Table 1. Manipulation Detection Rules

Categories	Description	Specific rules	# Manipulated reviews
Rules based on reviewer	Reviews with extreme ratings (all five stars) are highly likely to be fraud reviews (Heydari et al., 2016). The user who gives high (low) overall rating for all of the platforms on which they post reviews is suspicious to manipulate the reviews. Thus, we label all the reviews posted by this kind of reviewer as suspected manipulated.	User gives 20 of total rating for all of the platforms on which they post reviews.	110
		User gives 4 of total rating for all of the platforms on which they post reviews.	160
	The similarity of all reviews posted by the reviewer is above a certain threshold.	The similarity of all reviews posted by the reviewer is above 0.9	1876
	When a user releases review on multiple platforms, only one platform has lower overall rating.	Users' rating for platform above 16 is viewed as high rating, lower than 8 as low rating.	868
	Unusually active users are more likely to write abnormal reviews which are manipulated by the P2P lending platform; we take the reviews this user released on this certain platform as suspected	The number of reviews a user has wrote is above 3(95% quantile), with a low standard deviation 0.707(33% quantile) of total ratings in 21 days on a	2042

	manipulated reviews.	certain platform.	
		The text similarity of users' reviews on a certain platform all above 0.7	1367
Rules based on platform	The textual similarity of the platform's reviews is above a certain threshold.	The text similarities of a platform's reviews are all above 0.7	13
	Extreme large (small) mean and small (large) variance of all reviews' total rating on the platform.	If mean of a platform's rating above 16 and standard deviation lower than 3.83(60% quantile), or mean of platform's rating lower than 13 and standard deviation above than 3.83(70% quantile)	101
Rules based on review	If a review has high (low) total rating but the overall evaluation is non-recommended (recommend), we regard this comment as a suspected manipulated review.	If the overall evaluation of a review is recommend (not recommend) but the total rating is 4 (20).	3235
	Higher text similarity between two reviews, more likely both the review is manipulated (Zhao et al., 2013). So the second rule of reviews is: if text similarity between two reviews is greater than a certain threshold, both reviews are abnormal.	If the text similarity of the two reviews is greater than 0.9.	5117
Total			12201

Measurement

All the key variables used in the research model are listed in Table 2.

Table 2 Definition of Variables

	Variables	Description
Review helpfulness	*review_helfulness_ratio*	The ratio of helpful votes of a review.
manipulation	*is_abnormal*	1 if a review is manipulated and 0 otherwise
review readability	*average_word_length*	Average Chinese character of each Chinese word
	average_sentence_length	Average Chinese word count in each sentence
review depth	*review_word_count*	The number of total of Chinese words in a review
Platform quality	*shareholder_Soccontrolled_coded*	1 if state-owned enterprises hold shares in the platform and 0 otherwise
	registered_capital	Registered capital of the platform
	avg_volume	Average daily volume of turnover in the past year
Control variables	*is_recommend*	1 if a review is positive and 0 if a review is negative
	review_duration	The number of days between releasement and crawling.

Review manipulation and helpfulness

The dummy variable *is_abnormal* refers whether or not the review is manipulated. To get the dependent variable *review_helpfulness_ratio*, we divide its total number of votes by its total number of helpful votes for each review (Korfiatis et al., 2012).

$$review\_helpfulness\_ratio = \frac{helpful\ votes}{total\ votes}$$

Review quality

Text readability refers cognitive effort required for a person to comprehend a piece of text (Zakaluk and Samuels, 1988). The popular readability index like ARI (Automated Readability Index) is for a typical English language text (Senter and Smith, 1967).

$$ARI = 4.71 \times \left(\frac{characters}{words}\right) + 0.5 \times \left(\frac{words}{sentence}\right) - 21.43$$

As for readability of Chinese text, we calculate two variables defined as follows. *average_word_length* and *average_sentence_length* and *review_word_count* for each review (Chen et al., 2018). Reviews with small *average_word_length* and *average_sentence_length* are more readable. Besides, the larger word count means high text depth which contain more information.

$$average\_word\_length = \frac{total\ characters}{review\_word\_count}$$

$$average\_sentence\_length = \frac{review\_word\_count}{total\ sentences}$$

Firm quality

We use three proxy variables to measure the quality of the P2P lending platforms. *Avg_volume* is the average daily loan volume in the past year of the P2P lending platform. *Resitered_capital* is registered capital at the founding of the platform. *Shareholder_Soccontrolled_coded* is a dummy variable that shows whether the state-owned enterprises hold the shares of the platform.

Control variables

Our research model also includes two control variables. The first control variable *is_recommend* refers to the overall evaluation of the review indicating whether the reviewer would like to recommend the P2P lending platform to others. This variable is coded as 1 if it is positive and 0 if negative. To answer research question 4 and 5, the dataset is split as positive reviews and negative reviews. The second control variable *review_duration* refers to how long a review has been posted in the forum. The longer a review is presented, the more likely it is to get more helpful votes (Siering et al., 2018).

Descriptive Statistics

Descriptive statistics of the variables are listed in table 3, which shows that the suspected manipulated reviews account for about 9% of the total reviews and 60 percentages of the reviews get overall recommend evaluations (i.e., positive).

Table 3 Descriptive Statistics

Variables	Count	Mean	Std	Min	50%	Max
review_helfulness_ratio	88973	0.506763	0.454365	0	0.5	1

review_duration (Day)	88973	222.7072	182.4132	2	193	1383
registered_capital (10000 Yuan RMB)	88973	10006.18	12309.43	500	10000	250000
avg_volume (10000 Yuan RMB)	88973	4134.874	6391.153	5.1003	1295.425	368658
average_sentence_length	88973	16.85864	11.47258	0.0556	14	185
review_word_count	88973	11.69719	8.624103	0	9	128
average_word_length	88973	1.679393	0.20903	1	1.66667	21
	Dummy variables (percentage)					
	1	0				
Shareholder_Soccontrolled_coded	95.55%	0.45%				
is_abnormal	8.67%	91.33%				
is_recommend (positive or negative)	67.24%	32.76%				

Results

Descriptive evidence

Before answering the research questions using path analysis, we presented some relevant descriptive evidence of the manipulation of review helpfulness. The reviews are grouped by attitude (positive or negative) and manipulation (manipulated or normal). Thus, we get four groups shown in figure 3. Firstly, we depict the mean helpfulness for each group. As we can see, the positive manipulated reviews are most useful (mean ratio= 0.569), and the negative manipulated reviews are most useless (mean ratio = 0.485). So, we suspect that one possible reason is that P2P lending platform will upvote its manipulated positive reviews and downvote the fraud negative reviews it receives.

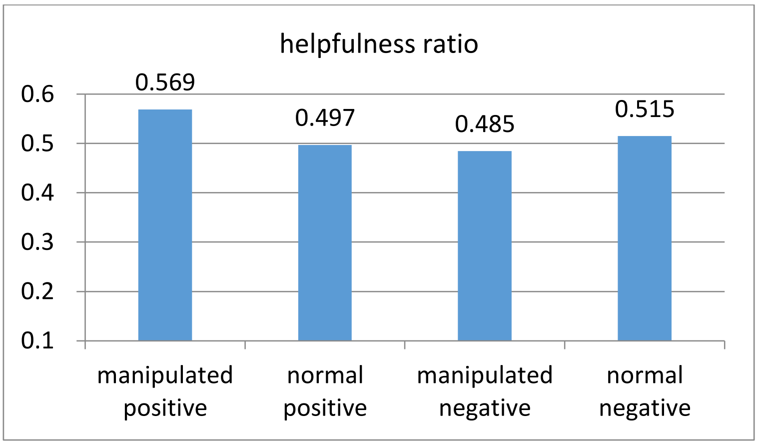

Figure 3 Helpfulness Ratios and Review Groups

To see if there is significant difference in review helpfulness ratio between groups, we perform ANOVA(One-way analysis of variance) on four groups (Fisher, 1918). As our data, helpfulness ratio, does not meet the homogeneity of variances assumption, so we run the Games Howell post hoc test to test the differences between groups (Ruxton and Beauchamp, 2008). As showed in the table 4, the helpfulness of manipulated positive reviews is significant greater than other three kinds. The manipulated negative reviews get the lowest helpfulness.

Table 4 Games Howell Post Hoc Test

(I) kind_cat	(J) kind_cat	Mean difference (I-J)	Standard error	Significant
manipulated positive	normal negative	0.054	.007	.000
	manipulated negative	0.084	.011	.000
	normal positive	0.072	.006	.000
normal negative	manipulated positive	-0.054	.007	.000
	manipulated negative	0.030	.010	.009
	normal positive	0.018	.003	.000
manipulated negative	manipulated positive	-0.084	.011	.000
	normal negative	-0.030	.010	.009
	normal positive	-0.013	.010	.544
normal positive	manipulated positive	-0.072	.006	.000
	normal negative	-0.018	.003	.000
	manipulated negative	0.013	.010	.544

Hypotheses testing

We adopt the path analysis to test the hypotheses. Path analysis was developed around 1918 by geneticist Sewall Wright, who wrote about it more extensively in the 1920s (Wright, 1921). It has since been applied to a vast array of complex modeling areas, including biology, psychology, sociology, and econometrics (Dodge, 2003). Path analysis can be viewed as a special case of structural equation modeling (SEM), in other words, path analysis is SEM with a structural model but no measurement model.

Mediating Role of Review Quality

As shown in Table 5, both average sentence length and review word count have a significant direct effect on review helpfulness ratio (*coefficient=-0.00186, p<0.01 and coefficient=0.00687, p<0.01*), while the effect of average word length is not significant, suggesting support H1. We find that manipulation (*is_abnormal)* has a significant positive direct effect on average word length and average sentence length (*coefficient=14.96, p<0.01 and coefficient=1.694, p<0.01*), but the direct effect is negative for review word count (*coefficient=-1.393, p<0.01*). This result suggests that a manipulated review tends to have longer sentence length and word length, but fewer words. So, H2 is supported. Table 6 suggests that the mediating effect on review helpfulness ratio in the path is significantly negative (*coefficient=-0.024428, p<0.01*), further indicates that manipulated review has lower review quality, leading to a lower review helpfulness ratio.

Table 5 Path Analysis Result

Hypothesis	Path	Direct
H1	average_sentence_length --> review_helpfulness_ratio	-0.00186*** (0.000170)
	average_word_length --> review_helpfulness_ratio	0.00764 (0.00729)
	review_word_count --> review_helpfulness_ratio	0.00687***(0.000225)
H2	is_abnormal --> average_sentence_length	14.96***(0.227)
	is_abnormal --> average_word_length	1.694***(0.0184)
	is_abnormal --> review_word_count	-1.393***(0.103)

Notes: standard error in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6 Mediating Role of Review Quality

Path	Mediator s	Indirect	Direct	Total
is_abnormal --> review_helpfulness_ ratio	average_sentence_length average_word_length review_word_count	-0.024428*** (0.0008335)	0.04002*** (0.005382)	0.015594*** (0.0054465)

Notes: standard error in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Manipulation of Review Helpfulness

After controlling the mediating effects of review quality, we will test the existence of direct manipulation of review helpfulness. The results in table 6 show that the direct effect of manipulation on helpfulness is positively significant (*coefficient=0.04, p<0.001*). So, H3 is supported if we run the model with complete dataset. However, if we spit the dataset as positive reviews and negative reviews and run the model separately, it turns out that manipulation has a positive effect on helpfulness as for positive review (*coefficient=0.0610, p<0.001*), while it has negative effect on helpfulness for negative review (*coefficient=-0.0232, p<0.001*) (see table 7). So, only the positive manipulated review is more likely to receive higher helpfulness due to the manipulation of review helpfulness. Thus, H3 is partially supported. Meanwhile, the firm may manipulate to reduce the negative review helpfulness.

Firm Quality and Reviews Manipulation

In order to answer research question 4 and 5, we divided the data set into two datasets, positive review and negative review. The results in table 7 indicate that high quality firms tend to manipulate more positive reviews (*coefficient=0.00000294, p<0.001*). So H4 is supported. It also suggest that high quality firms will receive more negative manipulated review (*coefficient=0.000000549,p<0.01*). Hence, H5 is supported.

Table 7 Path Analysis for Positive and Negative Reviews

Relationships	Positive reviews	Negative reviews
is_abnormal<—		
shareholder_soccontrolled_coded	0.00229 (0.00536)	0.0106 (0.00919)
registered_capital	-0.000000118 (0.000000102)	0.000000108 (0.000000120)
avg_volume	0.00000294*** (0.000000203)	0.000000549** (0.000000225)
average_sentence_length<—		
is_abnormal	14.71*** (0.262)	15.57*** (0.443)
review_helpfulness_ratio<—		
is_abnormal	0.0610*** (0.00654)	-0.0232** (0.00947)
average_sentence_length	-0.00307*** (0.000230)	-0.000117 (0.000245)
average_word_length	0.0132 (0.00948)	-0.00992 (0.0112)

review_word_count	0.00828*** (0.000322)	0.00500*** (0.000304)
review_duration	-0.000282*** (0.0000104)	-0.0000358*** (0.0000137)
average_word_length<—		
is_abnormal	1.697*** (0.0217)	1.686*** (0.0347)
review_word_count<—		
is_abnormal	-1.271*** (0.108)	-1.467*** (0.228)
N	59824	29149

Notes: standard error in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A summary of hypotheses testing was displayed in the table 8.

Table 8 Hypotheses Testing Summary

Hypotheses	Result
H1: Review quality (readability and words count) will improve its helpfulness	Supported
H2: A manipulated review tends to be of lower text quality.	Supported
H3: A manipulated review is more likely to receive higher helpfulness due to the manipulation of review helpfulness.	Partially supported
H4: high quality firms tend to manipulate more positive reviews.	Supported
H5: high quality firms will receive more negative manipulated review.	Supported

Conclusions

This paper examines the existence of review helpfulness manipulation and the relationship between firm quality and review manipulation. Using the archival data from the P2P lending forum in China, we conduct a path analysis to assess the manipulation's direct and indirect effects on helpfulness. We get three main findings as follows. First, because of helpfulness manipulation, a manipulated positive review is more likely to receive higher helpfulness, while a manipulated negative review is more likely to get lower helpfulness. Second, a manipulated review tends to be lower quality in terms of readability and word count, which could improve review helpfulness. Third, high quality firms tend to manipulate more positive reviews. In addition, high quality firms will receive more negative manipulated review from its competitors.

These findings shed lights on the areas of online review manipulation and P2P lending market. In addition, as helpfulness of review can be manipulated by firms, the usefulness data based on the number of users' votes may be not a valid measurement of the true helpfulness of online review. So, future research should develop alternative measure of the true helpfulness of online reviews, and investigate how much the helpfulness votes affecting subsequent consumers' decisions.

Acknowledgments

This work is supported by the Chinese Ministry of Education Humanities and Social Sciences Fund (18YJAZH142), and the Natural Science Foundation of China (71771159).

References

Chen, X., Dezhu, Y. and Jie, D. 2018. "Can readability of loan description promote lending success rate of online," *China Industrial Economics* (03), pp. 174-192.

Cheung, C. M. K., and Thadani, D. R. 2012. "The impact of electronic word-of-mouth communication: A literature analysis and integrative model," *Decision Support Systems* (54:1), pp. 461-470.

Dellarocas, C. 2006. "Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms," *Management Science* (52:10), pp. 1577–1593.

Dodge, Y. 2003. "The Oxford Dictionary of Statistical Terms," OUP. ISBN 0-19-920613-9.

Harmon, A. 2004. "Amazon Glitch Unmasks War of Reviewers," The New York Times (February 14, 2004).

Heydari, A., Tavakoli, M., and Salim, N. 2016. "Detection of fake opinions using time series," *Expert Systems with Applications* (58), pp. 83–92.

Hu, N., Bose, I., Gao, Y., and Liu, L. 2011. "Manipulation in digital word-of-mouth: a reality check for book reviews," *Decision Support Systems* (50:3), pp. 627-635.

Hu, N., Bose, I., Koh, N. S., and Liu, L. 2012. "Manipulation of online reviews: an analysis of ratings, readability, and sentiments," *Decision Support Systems* (52:3), pp. 674-684.

Hu, N., Liu, L., and Sambamurthy, V. 2011. "Fraud detection in online consumer reviews," *Decision Support Systems* (50:3), pp. 614-626.

Korfiatis, N., García-Barriocanal E., and S Sánchez--Alonso. 2012. "Evaluating content quality and helpfulness of online product reviews: the interplay of review helpfulness vs. review content," *Electronic Commerce Research and Applications* (11:3), pp. 205-217.

Lin, M. F., Prabhala, N. R., and Viswanathan, S. 2013. "Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending," *Management Science* (59:1), pp. 17-35.

Mayzlin, and Dina. 2006. "Promotional chat on the internet," *Marketing Science* (25:2), pp. 155-163.

Moffitt, K., and Burns, M. B. 2009. "What Does That Mean? Investigating Obfuscation and Readability Cues as Indicators of Deception in Fraudulent Financial Reports," in *AMCIS 2009 Proceedings*.

Mudambi, S. M., and Schuff, D. 2010. "What makes a helpful review? A study of customer reviews on Amazon. Com," *MIS quarterly* (34:1), pp. 185-200.

Ong, T., Mannino, M., and Gregg, D. 2013. "Product Reputation Manipulation: The Impact of Shill Reviews on Perceived Quality. PhD thesis.

Ronald, A. Fisher. 1918. "The Correlation Between Relatives on the Supposition of Mendelian Inheritance," Philosophical Transactions of the Royal Society of Edinburgh (52). pp. 399–433.

Ruxton, G.D., and Beauchamp, G. 2008. "Time for some a priori thinking about post hoc testing," Behavioral Ecology (19:3), pp. 690-693.

Senter, R. J., and Smith, E. A. 1967. "Automated readability index," retrieved from, http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0667273196.

Siering, M., Muntermann, J., and Rajagopalan, B. (2018). "Explaining and predicting online review helpfulness: The role of content and reviewer-related signals," *Decision Support Systems* (108), pp. 1–12.

Wang, Y. and Yu, C. 2015. "Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning," *International Journal of Information Management* (37:3), pp. 179-189.

Wright, S. 1921. "Correlation and causation, " *Journal of Agricultural Research* (20), pp. 557–585.

Zakaluk, B. L., and Samuels, S. J. 1988. "Readability: its past, present, and future," International Reading Association, 800 Barksdale Rd., PO Box 8139, Newark, DE l9714-8139.

Zhao, Y., Yang, S., Narayan, V., and Zhao, Y. 2013. "Modeling Consumer Learning from Online Product Reviews," *Marketing Science* (32:1), pp. 153–169.

Zhuang, M., Cui, G., and Peng, L. 2018. "Manufactured opinions: The effect of manipulating online product reviews," *Journal of Business Research* (87), pp. 24-35.