2-2011

Manipulation in digital word-of-mouth: A reality check for book reviews

Nan HU
Singapore Management University, nanhu@smu.edu.sg

Indranil BOSE
University of Hong Kong

Yunjun GAO
Zhejiang University

Ling LIU
University of Wisconsin - Eau Claire

Citation

HU, Nan; BOSE, Indranil; GAO, Yunjun; and LIU, Ling. Manipulation in digital word-of-mouth: A reality check for book reviews. (2011). *Decision Support Systems*. 50, (3), 627-635.

Available at: https://ink.library.smu.edu.sg/sis_research/8219

Decision Support Systems 50 (2011) 627–635

Contents lists available at ScienceDirect

Decision Support Systems

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

Manipulation in digital word-of-mouth: A reality check for book reviews

Nan Hu [a,b,*], Indranil Bose [c], Yunjun Gao [d,*], Ling Liu [a]

[a] Department of Accounting and Finance, University of Wisconsin Eau Claire, 105 Garfield Ave, Eau Claire, USA, 54701
[b] School of Information Systems, Singapore Management University, 80 Stamford Road, Singapore, 178902
[c] Department of Information Systems, School of Business, The University of Hong Kong, Pokfulam Road, Hong Kong
[d] College of Computer Science, Zhejiang University, Hangzhou, China, 310027

ARTICLE INFO

Available online 19 August 2010

Keywords:
Book reviews
Empirical study
Fraudulent manipulation
Online word-of-mouth
Regression analysis
Review management
Review manipulation

ABSTRACT

Built upon the discretionary accrual-based earnings management framework, our paper develops a discretionary manipulation proxy to study the management of online reviews. We reveal that fraudulent review manipulation is a serious problem for 1) non-bestseller books; 2) books whose reviews are classified as not very helpful; 3) books that experience greater variability in the helpfulness of their online reviews; and 4) popular books as well as high-priced books. We also show that review management decreases with the passage of time. Just like fraudulent earnings management, manipulated online reviews reflect inauthentic information from which consumers might derive wrong valuation especially for books with the above characteristics and be persuaded to purchase the wrong item. The findings from this research sound a note of caution for all consumers that make use of online reviews of books for making purchases and encourage them to delve deeper into the reviews without getting trapped in their fraudulent manipulation.

1. Introduction

Given that consumers are increasingly depending on user-generated content from non-traditional information channels, such as online reviews, to make purchase or investment decisions, this paper examines review management, defined as vendors, publishers, or writers consistently monitoring consumer online reviews, posting non-authentic messages to message boards, or writing inflated online reviews on behalf of customers when needed, with the goal of boosting their product sales, in the online review context.

Before the emergence of consumer-generated contents, consumers made their investment or purchase decisions based on information released through traditional channels, such as company financial reports. However, every now and then consumers might be mis-guided and make the wrong investment decisions if the financial reports they depended on were tampered with by the management of a firm who deliberately engaged in fraudulent earnings management in order to embellish the financial statements, meet a pre-specified target, and achieve better compensation. Hence the earnings management literature in accounting advocates that investors should look at the financial statements more carefully to detect a material and intentional misrepresentation of results. To gauge the information quality, previous literature has studied how to detect earnings management of traditional information, such as company's financial statements, using various methods, such as Jones' model [21] or a modified Jones' model [8,27].

However, we should be aware that, besides using traditional information, such as financial reports, customers are increasingly depending on other non-traditional information, such as online consumer-generated content, to make both stock valuation decisions and product purchase decisions. On the valuation side, accounting literature has documented that consumers act on user-generated content, such as messages on message boards, in order to make investment decisions. For example, Wysocki [30] examined whether daily changes in the volume of message board posting is correlated with changes in daily stock trading volume, absolute value of daily abnormal stock returns, and actual daily abnormal stock returns. Antweiler and Frank [1] found that investors view the information conveyed by stock message boards as valuable. They found that a positive shock to message board postings predicted negative returns on the next day. And based on message posts from Yahoo! Finance from April 2005 to April 2006, Gu et al. [17] showed that the weighted average sentiment of a stock message board could predict future abnormal stock returns, whereas the equal-weighted average sentiment of the message board had no such predictive power for future stock movements.

On the product purchase side, literature in economics and marketing has also suggested that consumers depend on online product reviews to make purchase decisions [2,4,29]. This literature provided useful insights by linking online reviews with sales. It showed a positive correlation between the average review score and product sales [3,4,10,15] or between the volume of reviews and sales

* Corresponding authors. N. Hu is to be contacted at Department of Accounting and Finance, University of Wisconsin Eau Claire, 105 Garfield Ave, Eau Claire, USA, 54701. Y. Gao, College of Computer Science, Zhejiang University, Hangzhou, China, 310027.
E-mail addresses: hun@uwec.edu, hunan@smu.edu.sg (N. Hu), bose@business.hku.hk (I. Bose), gaoyj@zju.edu.cn (Y. Gao), liul@uwec.edu (L. Liu).

0167-9236/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2010.08.013

[12,25]. Recent research has also showed that, besides the numeric rating, consumers also pay attention to other aspects of online consumer reviews to make purchase decisions, such as reviewer identity [13,20] and consumer sentiments incorporated inside textual comments [5,11,19].

Since consumer-generated content holds the power of influencing consumers' investment and purchase decisions, it is reasonable to assume that interested parties, such as publishers, vendors, or writers, will try to manipulate the final outcomes through posting non-authentic messages. To cope with review management, it is crucial and timely to understand the quality aspect of information released through this new digital channel. To date, there have been only a few analytical studies investigating the manipulation of online consumer reviews [9,26]. However, there are a reasonable number of cases that lead us to suspect that review manipulation is a serious problem in some industries. For example, the music industry is known to hire professional marketers to write favorable consumer opinions to promote the sales of new albums [26]. Publishers or authors of books boost ratings of their newly published books sometimes by paying someone or asking their friends or relatives to write glowing reviews. In one extreme case, Annalee Newitz[1] discussed how people manipulated online reputation systems like Digg. After noticing that a new story about a blog dedicated to showing photographs of crowds received enough 'diggs' to make it to the 'popular' list on the tech/design page, he conducted some investigation on how this happened and his finding revealed the following:

> I can tell you exactly how a pointless blog full of poorly written, incoherent commentary made it to the front page on Digg. I paid people to do it. What's more, my bought votes lured honest Diggers to vote for it too. All told, I wound up with a "popular" story that earned 124 diggs — more than half of them unpaid. I also had 29 (unpaid) comments, 12 of which were positive.I hired a Digg-gaming service called User/Submitter or U/S. This enterprise, run by one or more zealously anonymous individuals, advertises that it can help "submitters" get Digg stories noticed by paying "users" to digg them. There's a $20 sign-up fee and each digg costs $1, which gets split evenly between the service and the digger. U/S refunds money paid for any diggs the submitter doesn't get in a 48-hour period. I put down $450 for 430 diggs, but wound up getting refunded all but roughly $100 of that.

As we can see, it is neither rare nor difficult for a company, in this case a blogger, to pay someone to boost the perceived quality of a product or service. Feedback manipulation is not uncommon in the industry. For example, in March 2008 eBay announced that it would ban digitally delivered goods from either its auction or its listing services.[2] The reason was that the selling of electronically delivered items was used to manipulate a member's feedback ratings. eBay also announced that some of its members had become increasingly sophisticated in manipulating reputation on eBay and it vowed to fight such fraudulent practices aggressively. Feedback is the foundation of trust on eBay, and fraudulent purchasing, selling, or trading feedback undermined the trust within the eBay community.[3] According to recent statistics from NW3C/FBI 2007, online auction fraud accounted for 45% of the total cases referred to US law enforcement agencies and contributed to 33% of the total reported dollar loss.

Even though the above mentioned industry cases reveal the existence of online review manipulation, previous empirical literature that studied the impact of user-generated content on product sales implicitly assumes that user-generated content, such as online reviews, are written by actual previous customers, and not by publishers, vendors, or other self-interested parties. The goal of this paper is to come up with a discretionary review manipulation proxy that can empirically validate the above assumption based on publicly available data, and shed some light on online review manipulation by addressing the following research questions:

- Given that consumers and researchers have no access to the real identity of the reviewers, how can we show empirically that manipulation of online reviews of books exists?
- What is the effect of product characteristics, such as bestseller books versus non-bestseller books, high-priced books versus low-priced books, books receiving a high percentage of helpful reviews versus books receiving a limited percentage of helpful reviews, on review management?
- What is the effect of time on review management? For the same book, does online review manipulation lessen with the passage of time?

This paper proceeds as follows. Section 2 develops our discretionary review manipulation proxy and relates various attributes associated with online book sales such as popularity, price, popularity and variability in reviews, and time of sales to review manipulation in the form of several hypotheses and explains under what scenarios fraudulent review management is a serious problem. Section 3 presents the definition for all variables used in this research. Section 4.3 shows the results of the hypothesis testing as well as robustness checks for the findings. Section 5 discusses the findings and their implications, and provides some concluding remarks.

2. Discretionary review manipulation proxy and hypotheses development

In this section, we develop our discretionary manipulation proxy by building on the discretionary accrual-based earnings management framework. In the earnings management literature, the most commonly used vehicles for earnings management are either the choice of accounting procedures (methods) or the choice of estimates for a given accounting method. Since the effects of earnings management through changing accounting procedures are highly visible and can be easily 'undone' by outsiders, thereby defying the purpose of earnings management, it is more likely for managers to manage earnings through discretionary accruals [27].

Based on discretionary accruals, the accounting literature has used the following framework to uncover the existence of earnings management [8,27]:

$$DA_t = \beta_0 + \beta_1 Group_t + \sum_{i=1}^{K} \gamma_i X_i + \varepsilon_t \tag{1}$$

where

DA represents discretionary accruals;
Group is a dummy variable that is equal to one for firm-years during which earnings management is more likely to occur due to the stimulus identified by researchers. In other words, these firm-years share some type of common characteristics which researchers believe will drive the company to engage in earnings management.
X_i represents other relevant variables that may influence discretionary accruals.

The null hypothesis of no earnings management in response to the researcher's stimulus will be rejected if β_1, the coefficient of the *Group* variable, is significant and has the right hypothesized sign [8,27].

[1] http://www.wired.com/techbiz/people/news/2007/03/72832.
[2] http://blog.skipmcgrath.com/public/blog/200270.
[3] www2.ebay.com/aw/archive.shtml.

However, when β_1 is not statistically significant, the research is inconclusive as to whether no earnings management is present or the power of the test is too low because the earnings management test is always a joint test of a researcher's model assumptions (e.g., empirical proxy that the researchers need to come up with for estimation of *DA* that represents the unobservable discretionary accruals) and the earnings management hypotheses. Hence, coming up with the right discretionary accruals estimation (the dependent variable) is a key concern.

In order to come up with the right discretionary manipulation proxy in the online review manipulation context, we need to understand the nature and the motivation of review manipulation. First, we believe the goal of review management is to swing consumers' purchase decisions by posting inauthentic reviews. Hence, review management is not very different from slanting, which is defined as "the process of selecting details that are favorable or unfavorable to the subject being described" [18]. Slanting is commonly adopted by public relations, law, marketing, professional writing, and advertising to influence a third party's opinion. For example, Gentzkow and Shapiro [14] showed that instead of providing unbiased views, newspapers tailored their content extensively to fit the ideological predispositions of their readers with the goal of maximizing their readership. For the topics where reader beliefs diverged, newspapers segmented the market and slanted towards extreme positions [28]. The reason for newspapers slanting their news to cater to the preferences of their audiences is that readers tended to seek information that confirmed their beliefs [23].

Second, we believe that boosting rating, which implied giving a very positive rating to a product assuming the identity of a customer, is a common but also very feasible review manipulation strategy. In an online context, if potential customers know which reviews are posted by previous real customers and which reviews are written by authors, publishers, or any self-interested third parties, then those potential customers can undo the damage caused by these slanting reviews and filter such reviews out. Unfortunately, since all these slanting reviews are written either using a 'customer' name or assuming an anonymous identity, consumers cannot separate a slanting review from a positive review written by a zealous customer just by looking at the rating and the content of a review. Even manually inspecting the content of reviews cannot solve that problem because it is still very difficult to differentiate between the reviews unless some parts of a review are an identical duplicate of another review [6]. Since review manipulation by posting very positive comments under an assumed 'consumer' identity is less likely to be 'undone,' we expect that the effect of review manipulation will be reflected in the form of a ratings boost in many contexts.

Third, as in the case of earnings management where managers are more likely to engage in earnings management when facing poor prior earnings performance [7], we believe that it is more likely for publishers, vendors, or authors to engage in review manipulation when the product ratings are decreasing. In such a situation, publishers, vendors, or authors will write good reviews on behalf of actual customers to boost the ratings of their chosen books.

With the above three assumptions, we next develop our discretionary manipulation proxy. As we elaborated above, the goal of this manipulation is to increase the perceived product quality. Even though this manipulation may lead to an improved average consumer rating, such a strategy will inevitably lead to a high fluctuation of ratings over time. And such variability in ratings is a consequence of review manipulation that cannot be controlled by the manipulators. By incorporating this fluctuation of ratings and other measures, we come up with a discretionary manipulation proxy.

In the following section we explain why we believe there will be a high fluctuation in the overall ratings for items that are subjected to manipulation. Let us assume that a book received five consumer rating scores and these are 1, 2, 3, 4, and 5 respectively (average of rating scores = 3 and variance of rating scores = 2.5). The exact sequence of these five numbers as ratings is subject to the following two potential processes:

- The self-selection theory proposed by Li and Hitt [24] states that there are systematic differences in the ratings provided by early customers and late customers, and normally early stage adopters (customers) are more likely to leave a positive review. As a result, consumer reviews of a product received at the early stage are systematically positively biased. In such a case it is more likely for the five ratings considered in the above example to appear in a sequence like 5, 4, 3, 2, and 1. If we define the absolute difference between two adjacent rating as *Dif_Rating*, then the variance of *Dif_Rating* (*Var_Dif_Rating*) is 0 under such a circumstance;
- Review management driven by positive manipulation of ratings by vendors, publishers, authors, or any self-interested third parties in a book. In such a case it is more likely for the five ratings in the above example to appear in a sequence like 4, 3, 5, 2, and 1 or something similar. Under such a circumstance, the *Var_Dif _Rating* is 0.91. The reason 5 appearing in the middle is due to the third assumption that it is more likely for the potential manipulators to step in when facing a decrease in online ratings.

As we can see, if the reviews are driven by review management, the *Var_Dif _Rating* is 0.91 whereas that value is almost zero if the reviews are driven by self-selection. The above numerical example clearly demonstrates that for items with similar ratings, it is more likely that an item subjected to manipulation will show a higher value for *Var_Dif_Rating*. To conclude, we define our discretionary review manipulation proxy (DRMP) as the ratio of the variance of the absolute difference of two adjacent ratings (*Var_Dif_Rating*) and its variance of the ratings ($DRMP = Var\_Dif\_Rating_i \,/\, Var\_Rating_i$). We expect that a high likelihood of engaging in manipulation is associated with a high value of DRMP.

However, not all publishers, vendors, or authors engage in the practice of online consumer opinion manipulation and it is also likely that the above entities may take part in review manipulation for some types of products, but not others. The fear of reputation loss may prevent them from engaging in such activities. For example, the publishers, vendors, or authors of bestselling books may take into account the cost of reputation loss before deciding whether to engage in review manipulation or not. These entities normally have established their credibility in the business community and the user community over a long period of time. Hence, the cost of reputation loss for engaging in review manipulation will be much higher for publishers, vendors, or authors of bestselling books than for those associated with non-bestselling books. Furthermore, the gain achieved by publishers, vendors, or authors of bestselling books by taking part in such unethical behavior is also limited. Trading off the possible limited economic gain with the huge cost of reputation loss, we expect that it is less likely that vendors will engage in manipulation of online reviews for such bestselling books. So, our first hypothesis states that:

H1. Bestselling books serve as a non-manipulation indicator. In other words, it is less likely for publishers, vendors, or authors to engage in manipulating reviews for a bestseller book, and thus the reviews for such a book are associated with a smaller value of discretionary review manipulation.

For the same book, the manipulation of reviews may decrease with the passage of time. Again, we believe publishers, vendors, or authors conduct cost–benefit analyses before they decide whether to engage in review manipulation. As time elapses, the benefit of engaging in review manipulation decreases while the cost increases. The reason is that when a book is first available in the online market, there is few information channels that consumers can access to judge the quality or value of it, hence, consumers depend more on online reviews to make their purchase decision. Since at this early stage reviews have the highest impact on sales [20], publishers, vendors, or authors only need to post a few online reviews in order to sway consumers' purchase decisions. As time elapses, after a book receives a large number of consumer reviews, the cost of review manipulation becomes very high because the manipulation itself is not free and can be very time consuming. For example, on Amazon.com, a 'consumer' using a single IP address can only rate the same review once. For a book to receive a lot of reviews, vendors need to either switch to different PCs or use different proxies in order to post a decent number of biased reviews to counteract unfavorable reviews and make the manipulation work. Furthermore, at the later stage, besides online reviews, consumers have a number of alternative information channels, such as media coverage in newspapers, magazines, and television, or offline word-of-mouth, to gauge the quality of a book. Thus, we hypothesize that:

H2. After a book is released in the online market, the manipulation level will decrease with the passage of time leading to a smaller value of discretionary review manipulation.

When consumers read online reviews of books, they will pay attention to both the numerical score as well as the textual content of the reviews. High numeric ratings combined with rave and inauthentic content may convince potential consumers to buy the book at the first place. However, after reading the book, the consumers will find the true quality of the book and realize that they have been misled by the reviewers. If such a situation happens, consumers are likely to go back to the Web site and indicate that the review they read was not helpful. A lot of Web sites, including Amazon.com, implement such a mechanism to boost the quality of the consumer reviews they receive. Furthermore, many reviewers are working very hard by writing authentic consumer reviewers in order to climb up the ranks of Amazon.com's reviewers reward system. For example, the reviewer whose reviews are considered to be most helpful may be listed as a Top 1000 reviewer and will be assigned a Top 1000 badge whenever he or she posts a review on Amazon.com. A reviewer's prestigious position at Amazon.com is highly respected by the readers' community, and his/her reviews are treated as very helpful. Whenever such reviewers express their opinion about a book, potential buyers of books begin to take immediate notice. Hence, publishers, vendors, or authors will try to avoid manipulating user feedback for those books whose majority of reviews are viewed as very useful because the potential manipulators know that given there is less uncertainty for the quality of the books with high quality of reviews, the market is less likely to be influenced by their manipulation. If on an average, among all the reviews that a book receives, the reviews are considered to be very helpful, it signals that the quality of the reviews is high. Thus, we hypothesize that:

H3. Books whose reviews on an average are rated as highly helpful can serve as a non-manipulation indicator. In other words, it is less likely for publishers, vendors, or authors to engage in manipulating reviews for books that on average receive a higher percentage of helpful votes for their reviews. Thus such books enjoy a smaller value of discretionary review manipulation.

If, there is a large variability in the helpfulness of the reviews a book received, then this signals that there may not be a consensus about the online reviews quality that book received. Thus, we hypothesize that:

H4. The variance of the helpfulness of the reviews received by a book can serve as a manipulation indicator. In other words, it is less likely for publishers, vendors, or authors to engage in manipulating reviews of a book that experiences smaller variance in the helpfulness of its' reviews. Thus such a book will experience a smaller value of discretionary review manipulation proxy of its ratings.

In the accounting literature, Kim and Verrecchia [22] argue that in the case of anticipated events such as earnings announcements, investors are motivated to acquire private information because informed investors benefit more from private information than from less informed investors. Furthermore, based on Grossman's theory of information and competitive price systems [16], investors of large firms are more likely to engage in private information acquisition because large firms have greater share liquidity that hides informed trade, and thereby increases the return for private information acquisition. Following the same line of argument, we expect that publishers, vendors, or authors have a higher incentive to engage in online review manipulation for popular books because 1) such manipulation may be hidden by the large volume of online reviews, making it hard for the real customers to detect review management; and 2) they can make more financial gains from such a manipulation.

H5. Product popularity can serve as a manipulation indicator. In other words, it is more likely for publishers, vendors, or authors to engage in manipulating online reviews for popular books.

Following the same economic return argument as stated before, we also expect that:

H6. Price can serve as a manipulation indicator. In other words, it is more likely for publishers, vendors, or authors to engage in manipulating online reviews for books with a higher selling price.

3. Methods and measurement

3.1. Data

We collected the data used in this research from Amazon.com's Web Service in July 2005. A panel of books was randomly chosen in July 2005. For each item, we collected its price, sales, and reviews-related data. Our data included some very popular books, such as *The World Is Flat: A Brief History of the Twenty-first Century* by Thomas L. Friedman, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* by Steven D. Levitt, among others. We further chose those books that had received more than 100 consumer ratings[4] to make sure that these books have been in the market long enough to go through the entire manipulation lifecycle, from high to low to negligible. The collected data pertains to 1851 books.

3.2. Definition of variables

In this section, we discuss how we measure the different variables used in this research.

3.2.1. Bestseller dummy

To determine whether a book is to be considered a bestseller, we obtained the bestseller book list proposed by *The New York Times*. Then we compared the list of books obtained from Amazon.com with *The New York Times* Bestseller List. Since *The New York Times* Bestseller List changed over time, we made sure that the bestseller list we used was within the same period as the data collected from Amazon.com when we did the comparison between the two lists. This resulted in the identification of 77 bestseller books out of 1851 books. The variable *Bestseller_Dummy* is equal to 1 for those 77 books and 0 for other books.

3.2.2. Time dummy

To test whether there will be less manipulation of reviews after a book is listed on Amazon.com as time elapsed, we divided the reviews

[4] Changing the cut-off point to another number, such as 130, 120, 110, etc., did not change the results qualitatively.

of each book into two batches: batch 1 included the first 50% reviews received by a book; and batch 2 included the second 50% reviews that were received by the same book. As we explained in Section 2, the manipulation cost increased with the passage of time. Therefore, batch 1 represented the time period in which manipulations are likely to occur more frequently; whereas batch 2 represented the time period when manipulation is less likely. The variable *Time_Dummy* is equal to 1 for reviews belonging to batch 2 (late batch); and 0 otherwise

3.2.3. Average helpful vote dummy and variance of helpful vote dummy

For each review, we divided its total number of votes by its total number of helpful votes to get its helpful vote ratio. Then for each book and for each time batch, we estimated its average helpful vote ratio and its variance of helpful vote ratio. Subsequently, we defined high-average-helpful vote dummy (defined below), and high-variance-helpful vote dummy (defined below) separately for each batch respectively. Comparing the discretionary review manipulation proxy of these two batches of reviews helped us uncover the existence of review manipulation.

Next, we provide the detailed definition for each variable used in this study. Assume that book *j* received *n* reviews for each batch of reviews. So the total number of reviews received for book *j* is *2n*. For each review that is posted, Amazon.com also revealed how many customers read that review (*Total_Vote*) and how many considered it to be 'useful' (*Helpful_Vote*). We used these two variables with respect to a review *i* to define the average of helpful votes ($Avg\_Helpful\_Vote_j$) and variance of helpful votes ($Var\_Helpful\_Vote_j$) for book *j* as shown below:

Variance of Dif_Rating for book j

$$Var\_Dif\_Rating_j = Variance\left(abs\left(Rating_{t+1,j} - Rating_{t,j}\right)\right)$$

Average helpful vote for book *j*

$$Avg\_Helpful\_Vote_j = Average\left(\sum_{i=1}^{n_i} \frac{Helpful\_Vote_{i,j}}{Total\_Vote_{i,j}}\right)$$

Variance of helpful vote for book *j*

$$Var\_Helpful\_Vote_j = Variance\left(\frac{Helpful\_Vote_{i,j}}{Total\_Vote_{i,j}}\right)$$

3.2.4. High-average-helpful vote dummy

Based on the average number of helpful votes, we classified a book into a high-average-helpful or a low-average-helpful sub-group. The high-average-helpful group consists of those books whose average of the helpful vote for reviews is greater than or equal to the median of the helpful vote for that batch, whereas the low-average-helpful group includes those books whose average of the helpful vote for reviews is below the median. Thus, *High_Avg_Helpful_Vote_Dummy* is equal to 1 for the books included in the high-average-helpful group; otherwise, it takes a value of 0. Recall that we have an early stage reviews batch (consisting of the first 50% reviews) and a late stage reviews batch (consisting of the second 50% reviews). For each book, we defined its *High_Average_Helpful_Dummy* for its early and late batches respectively.

3.2.5. High-variance-helpful vote dummy

Based on the variance of the helpful votes received, we classified a book into a high-variance-helpful or a low-variance-helpful sub-group. The high-variance-helpful group consists of those books whose variance of helpful vote is greater than or equal to the median variance of helpful vote scores for that batch, while the low-variance-helpful group includes those books whose variance of helpful vote score is below the median. *High_Var_Helpful_Vote_Dummy* is equal to 1 for books that are included in the high-variance-helpful group; and is equal to 0 for all other cases. Again, for each book, we defined its *High_Var_Helpful_Vote_Dummy* for its early and late batches respectively.

3.2.6. Popularity

Sales of a product can be a proxy for its popularity. Instead of providing the actual sales number, Amazon.com provides the sales rank information of a listed item. The product sales rank is shown in descending order where 1 represents the best selling product. Consequently, there is a negative correlation between product sales and sales rank. We used *Log (SalesRank)* as a proxy for popularity where a high value of the *SalesRank* variable indicated low popularity.[5]

Even though on the surface it may seem that 'bestseller' and 'salesrank' are both measurements of an item's popularity, they do not necessarily measure the same thing because bestselling books selected by *The New York Times* may not enjoy very high sales on Amazon.com. In fact, there are many reasons that may lead a particular book to gain popularity in the Amazon.com community, regardless of whether it has been selected by *The New York Times* as the bestseller or not. At the same time, there is no guarantee that a bestseller book listed by the New York Times will be a popular book on Amazon.com. We checked the sales rank of *The New York Times* bestselling books on Amazon.com, and found that more than 10% of the bestsellers have sales rank greater than 100,000 in Amazon and some bestsellers even have sales rank greater than 200,000. This supported our conjecture that it is not necessary for bestsellers listed on *The New York Times* to be popular books on Amazon.com.

3.2.7. Price

As explained before, we used price to capture the financial incentive of manipulators because vendors will have a higher return for a high-priced product if they are able to manipulate the final outcome.

3.3. Empirical model

Recall that the accounting literature has used accruals-based tests to document earnings management using a linear framework [7,21]. Following that line of work, we linked the discretionary review management proxy to different dummy variables representing the situation where we believed that review management was more likely to occur due to the stimulus we identified above. Our final two models are as shown below:

Model 1.

$$\begin{aligned}\frac{Var\_Dif\_Rating_{j,m}}{Var\_Rating_{j,m}} = {} & \beta_0 + \beta_1 Bestseller\_Dummy_{j,m} + \beta_2 Time\_Dummy_{j,m} \\ & + \beta_3 Avg\_Helpful\_Vote\_Dummy_{j,m} \\ & + \beta_4 Var\_Helpful\_Vote\_Dummy_{j,m} \\ & + \beta_5 Log\left(Salesrank_{j,m}\right) + \beta_6 Log\left(Price_{j,m}\right) + \varepsilon_{j,m}\end{aligned}$$

Model 2.

$$\begin{aligned}\frac{Var\_Dif\_Rating_{j,m}}{Var\_Rating_{j,m}} = {} & \beta_0 + \beta_1 Bestseller\_Dummy_{j,m} + \beta_2 Time\_Dummy_{j,m} \\ & + \beta_3 Avg\_Helpful\_Vote_{j,m} + \beta_4 Var\_Helpful\_Vote_{j,m} \\ & + \beta_5 Log\left(Salesrank_{j,m}\right) + \beta_6 Log\left(Price_{j,m}\right) + \varepsilon_{j,m}\end{aligned}$$

[5] Using sales rank instead of sales has been a commonly approach in this line of work (e.g. 3, 4, 9, 13).

where $m = 1$ or 2 represents batch 1 (early batch) and batch 2 (late batch) respectively, and j stands for the jth product.

The only difference between Model 1 and Model 2 is that we converted the continuous variables representing average of helpful votes and variance of helpful votes in Model 2 to corresponding dummy variables in Model 1 because discretionary accrual framework requires the variable representing the situation where manipulation is more likely to happen to be a dummy variable. Recall that in Section 2 we hypothesized that vendors had higher incentive to engage in manipulation of online consumer reviews for the books that met one of the following criteria:

- Books that are non-bestsellers; or
- The same book but at the early stage
- Books whose reviews on average receive a lower percentage of helpful votes; or
- Books that show higher variances in the scores representing the helpfulness of their reviews; or
- Popular books; or
- High-priced books

The hypotheses related to manipulation of reviews will be supported if β_1 or β_2 or β_3 or β_5 is significantly less than zero; while β_4 or β_6 is significantly greater than zero.

4. Results and discussion

Table 1 shows the summary statistics of our collected data. In our sample, we had about 4% bestselling books, and on average only 39% of the customers thought the reviews on Amazon.com were useful. This indicated that many reviews might be manipulated. Furthermore, the mean of the variance of the helpful votes for reviews was small and only about 0.065. This combined with the low value of average helpful votes for reviews (0.3931) meant that there was relatively homogeneous belief among customers about the non-helpfulness of online consumer reviews.

4.1. Are publishers, vendors, or writers more likely to manipulate non-bestseller books?

To study whether publishers, vendors, or writers are consistently monitoring the online review channel, and will step in to manipulate online reviews, we estimated Model 1 using Ordinary Least Squares (OLS) and reported the results in Table 2. The coefficient for the *BestSeller_Dummy* variable was negative and significant at 10% level of significance (Para = −0.0525 and t-stat = −1.83), revealing that bestselling books did enjoy a smaller discretionary review manipulation. This proved that publishers, vendors, or writers did engage in online review manipulation but it was less likely for them to manipulate the consumer reviews of the bestselling books because 1) they were more concerned about their own loss of reputation when dealing with bestselling books; 2) these types of books already enjoyed very good sales; and 3) normally there was already good media coverage about bestselling books. In such a case, consumers had many different channels to assess the quality of a bestseller. Hence, it was hard as well as costly for vendors to try to influence outcomes of reviews for a bestseller through manipulation. This finding supported our first hypothesis.

Table 2
Estimation of the review manipulation.

$$\frac{Var\_Dif\_Rating_{i,m}}{Var\_Rating_{i,m}} = \beta_0 + \beta_1 Bestseller\_Dummy_i + \beta_2 Time\_Dummy_{i,m}$$
$$\beta_3 Avg\_Helpful\_Vote\_Dummy_{i,m}$$
$$+ \beta_4 Var\_Helpful\_Vote\_Dummy_{i,m}$$
$$\beta_5 Log(Salesrank_i) + \beta_6 Log(Price_i) + \varepsilon_i$$

Variables	Coefficients (t-stat)
Intercept	1.2407***
	(32.46)
Bestseller_Dummy	−0.0525*
	(−1.83)
Time_Dummy	−0.0619***
	(−5.42)
High_Avg_Helpful_Vote_Dummy	−0.1255***
	(−10.42)
High_Var_Helpful_Vote_Dummy	0.0109
	−0.94
Log(Salesrank)	−0.0161***
	(−6.23)
Log(Price)	0.0286**
	(2.37)
N	3702
Adj. R^2	0.0402

Notes: 1) The variable definitions are provided in Section 3.2. 2) All p-values are based on two-tailed tests.* indicates 10% level of significance; ** indicates 5% level of significance; and *** indicates 1% level of significance. 3) Salesrank and Price variables are log transformed. Salesrank is shown in descending order and 1 represents the bestselling book.

4.2. Does manipulation decrease over time?

In this sub-section, we examine the temporal effects of review manipulation (Hypothesis 2). As seen in Table 2, the coefficient for *Time_Dummy* was negative and significant at 1% level of significance (Para = −0.0619 and t-stat = −5.42), indicating that right after an item was listed on Amazon.com, publishers, vendors, or writers actively engaged in online review manipulation. However, as time went by and as books received more authentic consumer reviews, these entities had less incentive to conduct review manipulation because it was less likely for them to influence readers when a sufficient number of authentic consumer reviews existed.

4.3. Helpfulness of reviews and manipulation

The coefficient before the *High_Avg_Helpful_Vote_Dummy* variable represented the additional discretionary review manipulation

Table 1
Summary statistics.

Variables	Minimum	Q1	Mean	Median	Q3	Std. dev.
Bestseller_Dummy	0.0000	0.0000	0.0412	0.0000	0.0000	0.1987
Var_Dif_Rating	0.0000	0.7118	1.2708	1.2118	1.7290	0.7453
Var_Rating	0.0000	0.6531	1.3170	1.2092	1.8878	0.8370
Avg_Helpful_Vote	0.0000	0.2370	0.3931	0.4253	0.5434	0.2006
Var_Helpful_Vote	0.0000	0.0407	0.0650	0.0633	0.0846	0.0879
Salesrank	4	2291	59,111	11,811	56,471	122,410
Price	2.4000	7.9900	13.5576	11.1000	16.2000	9.5295

Notes: the variable definitions are provided in Section 3.2.

associated with books that received a high percentage of helpful votes. This coefficient was significantly less than zero at the 1% level of significance (Para = −0.1255 and t-stat = −10.42). At the same time, from Table 2, we observed that the coefficient before the *High_Var_Helpful_Vote_Dummy* was not significantly different from zero (Para = 0.0109 and t-stat = 0.94). Overall, it indicated that for books whose online reviews were considered as less helpful, there was a higher chance that the reviews of these books might be inauthentic (supporting Hypothesis 3). We did not find evidence showing that the high variance in the helpfulness score for online reviews received by a book gave rise to the possibility that the reviews might be tampered with (failed to support Hypothesis 4). However, when we replaced the high-variance and high-average-helpful dummy variables in Model 2 with continuous variables, we found partial support for Hypothesis 4 (Table 3) at the 5% level of significance.

4.4. Popularity, price, and manipulation

We found evidence that it was more likely for publishers, vendors, or writers to engage in online review manipulation for popular as well as high-priced books (supporting Hypotheses 5 and 6). In Table 2, the coefficient for the *Log(Salesrank)* variable was significantly less than zero (Para = −0.0161 and t-stat = −6.23) at the 1% level of significance, whereas the coefficient for the *Log(Price)* variable was significantly greater than zero (Para = 0.0286 and t-stat = 2.37) at the 5% level of significance.

4.5. Check on robustness of results

In this section, following the practice adopted in earnings management literature, we used Model 2 in which the average of helpful vote scores and the variance of helpful vote scores were converted to continuous variables to test Hypotheses 1 and 2 only. Recall that according to discretionary accruals, a researcher needed to use a dummy variable to divide the companies into two groups, one that researchers believed was more likely to engage in earnings management, while the other that was less likely to do so. The rest of the variables in the regression model were included as control variables.

From Table 3 we observed that the coefficients of both the bestseller dummy variable and the time dummy variable were significantly negative, showing that it was less likely for publishers, vendors, or authors to engage in online review manipulation for bestsellers and the manipulation decreased with elapsed time.

Table 3
Estimation of the review manipulation.

$$\frac{Var\_Dif\_Rating_{i,m}}{Var\_Rating_{i,m}} = \beta_0 + \beta_1 Bestseller\_Dummy_i + \beta_2 Time\_Dummy_{i,m}$$
$$\beta_3 Avg\_Helpful\_Vote_{i,m} + \beta_4 Var\_Helpful\_Vote_{i,m}$$
$$\beta_5 Log(Salesrank_i) + \beta_6 Log(Price_i) + \varepsilon_i$$

Variables	Coefficients (t-stat)
Intercept	1.3042***
	(33.46)
Bestseller_Dummy	−0.0604**
	(−2.1)
Time_Dummy	−0.0581***
	(−5.1)
Avg_Helpful_Vote	−0.3305***
	(−10.9)
Var_Helpful_Vote	0.1473**
	−2.26
Log(Salesrank)	−0.0168***
	(−6.5)
Log(Price)	0.0306**
	(2.53)
N	3702
Adj. R^2	0.0424

Notes: 1) The variable definitions are provided in Section 3.2. 2) All p-values are based on two-tailed tests. * indicates 10% level of significance; ** indicates 5% level of significance; and *** indicates 1% level of significance. 3) Salesrank and Price variables are log transformed. Sales rank is shown in descending order and 1 represents the bestselling book.

5. Discussions, conclusions, and future research

The advances in technology and the widespread use of the Internet has opened up many new channels, such as consumer reviews, message boards, and blogs, for consumers to exchange their opinions about products sold and services provided by a firm. These channels combine the immediacy of up-to-the-minute posts, normally latest first, to convey consumers' passions and point of view about a product or a firm. Due to their increasingly influential power over consumer purchase and valuation decisions, we utilized online book reviews to test the authenticity of the reviews and to see if there was any kind of fraudulent activities taking place.

In this study, we used data collected from Amazon.com to document that publishers, vendors, or authors consistently manipulated online consumer reviews. If a firm decided to use manipulation techniques, it was more likely for the firm to manipulate the reviews of 1) non-bestselling books; 2) books whose reviews were classified as not very helpful; 3) books with a greater divergence in the helpfulness of their online reviews; 4) popular books; and 5) high-priced books.

In comparison to earnings management which is widely studied in accounting, we believe that review manipulation is a more serious problem because:

1) Very limited theoretical support exists for review manipulation. Unlike the rich literature in earnings management, review manipulation is a new phenomenon that has recently caught the attention of both academic researchers as well as industry practitioners. Thus, at present, there is very limited understanding about how to detect review manipulation and under what circumstances it was a serious issue.
2) There exists very limited publicly available data that can be used to study this phenomenon. In sharp contrast, research on earnings management can use publicly released financial reports to uncover the fraudulent practices adopted by the firms. For example, publicly traded companies are required to file annual reports to disclose their financial status. So, based on the reported accrual numbers over time, researchers are able to develop various mechanisms to detect abnormal accruals to uncover the practice of earnings manipulation. However, very little data that can help uncover manipulation of reviews is publicly available. One might suspect that if one consumer kept using the same IP address to post very good ratings on specific items across multiple categories at the same time while spending almost no time in reading others' comments, this might be an indication of manipulation. However, all these detailed time-stamped data are only available to companies that own the e-commerce Web sites. Normally these companies are very reluctant to share such sensitive data with the public.
3) Even manually inspecting the reviews cannot guarantee that we can detect the manipulated reviews because it is still very difficult to differentiate between them [6].

The biggest challenge for this line of work is that, without having access to the proprietary data of e-commerce companies, researchers need to come up with a discretionary review manipulation proxy based on only publicly available data. We believe that the discretionary review manipulation idea that we came up with is a good starting point. Such a manipulation proxy is very important for policy makers, e-commerce providers, and end consumers. Without forcing companies to release their proprietary data, which might violate the consumers' privacy concerns, policy makers can use our method to check the existence of review management or blog opinion management for different product categories and for different communities in order to determine whether it is necessary for the government to regulate the communications within these online communities. For example, should the Securities and Exchange Commission force each blogger or message board poster to release their true identity when they post new messages since many such messages may have been tampered with and may have caused herding behavior of investors? For e-commerce companies, our idea of discretionary review manipulation can help them uncover different signals (variables) representing where and when manipulation is more likely to happen. And for end consumers, they can apply our framework to determine if the characteristics of the product they intend to buy matches with those for which review manipulation takes place and adjust their perception about the quality of the product before making a purchase decision.

To the best of our knowledge, this is the first empirical paper that has developed the idea of discretionary review manipulation proxy and has revealed the existence of online review manipulation for books. Our results showed that consumers should be more cautious in interpreting information from reviews for non-bestselling books, popular and high-priced books, and books with higher divergence in the helpfulness of their scores for online reviews, because vendors have a high incentive for conducting review manipulation for the above categories of books.

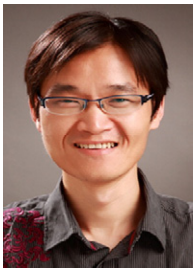
This paper is just the beginning of a long journey. It only sheds light on how to detect review manipulation. To really win the war against untruthful and misleading reviews, we urge future researchers to come up with better review manipulation proxy using various methods, such as pattern recognition, text mining, etc. We should also remind researchers again that one limitation of detecting review management is that, like market efficient hypothesis test, review management test results depend on the model that the researcher is using. So, it is always a joint test of review management and the review manipulation model used. However, despite the existence of the joint-test problem, tests of review management improve our understanding of the behavior of the review manipulator over time and across different categories. It helps uncover the practices of the online review market. If possible, future research should check its model assumptions with the back-end data from the field before drawing more insightful conclusions.

References

[1] W. Antweiler, M. Frank, Is all that talk just noise? The information content of internet stock message boards, The Journal of Finance 59 (2004) 1259–1294.
[2] P. Chatterjee, Online reviews: do consumers use them? Advances in Consumer Research 28 (1) (2001) 129–133.
[3] J. Chevalier, A. Goolsbee, Measuring prices and price competition online: Amazon and Barnes and Noble, Quantitative Marketing and Economics 1 (2) (2003) 203–222.
[4] J. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, Journal of Marketing Research 43 (3) (2006) 345–354.
[5] S.R. Das, M. Chen, Yahoo! For Amazon: sentiment extraction from small talk on the web, Management Science 53 (9) (2007) 1375–1388.
[6] S. David, T. Pinch, Six Degrees of Reputation: The Use and Abuse of Online Review and Recommendation Systems, Working Paper, 2005.
[7] L. DeAngelo, Managerial competition, information costs, and corporate governance: the use of accounting performance measures in proxy contests, Journal of Accounting and Economics 10 (1) (January 1988) 3–36.
[8] P.W. Dechow, R.G. Sloan, A.P. Sweeney, Detecting earnings management, The Accounting Review 70 (2) (April 1995) 193–225.
[9] C. Dellarocas, Strategic manipulation of internet opinion forums: implications for consumers and firms, Management Science 52 (10) (October 2006) 1577–1593.
[10] C. Dellarocas, N. Azad, M. Zhang, Using Online Ratings as a Proxy of Word-of-mouth in Motion Picture Revenue Forecasting, Working paper, 2005.
[11] V. Dhar, E. Chang, Does Chatter Matter? The Impact of User-generated Content on Music Sales, CeDER Working Paper 07-06, New York University, 2008.
[12] W. Duan, B. Gu, A.B. Whinston, The dynamics of online word-of-mouth and product sales — an empirical investigation of the movie industry, Journal of Retailing 84 (2) (June 2008) 233–242.
[13] C. Forman, A. Ghose, B. Wiesenfeld, Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, Information Systems Research 19 (3) (September 2008) 291–313.
[14] M.A. Gentzkow, J.M. Shapiro, Media Bias and Reputation, , September 14, 2005 Available at SSRN: http://ssrn.com/abstract=642362.
[15] D. Godes, D. Mayzlin, Using online conversations to study word of mouth communication, Marketing Science 23 (4) (2004) 545–560.
[16] S. Grossman, On the efficiency of competitive stock markets where traders have diverse information, Journal of Finance 31 (2) (May 1976) 573–585.
[17] B. Gu, P. Konana, A. Liu, B. Rajagopalan, J. Ghosh, Predictive Value of Stock Message Board Sentiments, Working Paper, University of Texas at Austin, 2006.
[18] S.I. Hayakawa, Language in Thought and Action, 5th Edition (Harcourt, Brace & Company, New York, 1990.
[19] N. Hu, N.S. Koh, V. Sambamurthy, The Value Implication of Online Consumer Review, American Accounting Association Seventeenth Annual Research Workshop on Artificial Intelligence and Emerging Technologies (AIET) in Accounting, Auditing and Tax (Anaheim/Orange, California, USA, 2008), 2008.
[20] N. Hu, L. Liu, J. Zhang, Do online reviews affect product sales? The role of reviewer characteristics and temporal effects, Information Technology and Management 9 (3) (2008) 201–214.
[21] J. Jones, Earnings management during import relief investigations, Journal of Accounting Research 29 (1991) 193–228.
[22] O. Kim, R. Verrecchia, Pre-announcement and event-period private information, Journal of Accounting and Economics 24 (1997) 395–419.
[23] J. Klayman, Varieties of confirmation bias, in: J. Busemeyer, R. Hastie, D. Medin (Eds.), Decision Making from a Cognitive Perspective: The Psychology of Learning and Motivation, 32, Academic Press, San Diego, CA, USA, 1995, pp. 365–418.
[24] X. Li, L.M. Hitt, Self selection and information role of online product reviews, Information Systems Research 19 (4) (2008) 456–474.
[25] Y. Liu, Word-of-mouth for movies: its dynamics and impact on box office receipts? Journal of Marketing 70 (3) (2006) 74–89.
[26] D. Mayzlin, Promotional chat on the internet, Marketing Science 25 (2) (2006) 157–165.
[27] M. McNichols, G.P. Wilson, Evidence of earnings management from the provisions for bad debts, Journal of Accounting Research 26 (1988) 1–31 Supplement.
[28] S. Mullainathan, A. Shleifer, The market for news, The American Economic Review 95 (4) (September 2005) 1031–1053.
[29] D. Reinstein, C. Snyder, The influence of expert reviews on consumer demand for experience goods: a case study of movie critics, Journal of Industrial Economics 53 (1) (2005) 27–51.
[30] P.D. Wysocki, Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards, Working Paper, University of Michigan, 1999.

Nan Hu is an Assistant Professor of Accounting and Finance at the University of Wisconsin at Eau Claire. He is also an Assistant Professor of Information Systems at Singapore Management University. He received his Ph.D. from the University of Texas at Dallas. Nan's research focuses on investigating the value implications and market efficiency of both traditional information (e.g. company financial report, analyst forecast, corporate governance, etc) and non-traditional information (e.g. blog opinion, online consumer reviews, etc), using a combination of theories from accounting, finance, marketing, information economics, sociology, psychology, and computer science. Nan's research has appeared at *JMIS (Journal of Management Information Systems), CACM (Communications of the ACM), JCS (Journal of Computer Security)*, and *IT&M (Information Technology and Management)*.

Indranil Bose is an associate professor at the School of Business, The University of Hong Kong. He holds a B. Tech. from the Indian Institute of Technology, MS from the University of Iowa, MS and Ph.D. from Purdue University. His research interests are in telecommunications, data mining, information security, and supply chain management. His publications have appeared in Communications of the ACM, Communications of AIS, Computers and Operations Research, Decision Support Systems, Ergonomics, European Journal of Operational Research, Information and Management, Journal of Organizational Computing and Electronic Commerce, Journal of the American Society for Information Science and Technology, Operations Research Letters etc. He is listed in the International Who's Who of Professionals 2005–2006, Marquis Who's Who in the World 2006, Marquis Who's Who in Asia 2007, Marquis Who's Who in Science and Engineering 2007, and Marquis Who's Who of Emerging Leaders 2007. He serves on the editorial board of Information and Management, Communications of AIS, and several other IS journals.



Yunjun Gao is an Associate Professor at the College of Computer Science, Zhejiang University, China. He received the PhD degree in computer science from Zhejiang University, China in 2008. Before joining Zhejiang University in 2010, he was a postdoctoral fellow at the School of Information Systems, Singapore Management University, Singapore during 2008–2010. His research interests include spatial databases, spatio-temporal databases, mobile and pervasive computing, and geographic information systems. He is a member of the ACM, ACM SIGMOD, IEEE, and CCF.

Ling Liu is an Assistant Professor of Accounting and Finance at the University of Wisconsin at Eau Claire. She received her Ph.D. in Accounting from the University of Texas at Dallas. Her research focuses on market efficiency, corporate governance, and relative performance evaluation.