Fraud detection in online consumer reviews ☆

Nan Hu [a,b,*], Ling Liu [a], Vallabh Sambamurthy [c]

[a] Department of Accounting and Finance, University of Wisconsin Eau Claire, 105 Garfield Ave, Eau Claire, United States
[b] School of Information Systems, Singapore Management University, 80 Stamford Road, Singapore
[c] Center for Leadership of the Digital Enterprise Eli Broad, College of Business, Michigan State University, East Lansing, MI 48824-1122, United States

ARTICLE INFO

ABSTRACT

Increasingly, consumers depend on social information channels, such as user-posted online reviews, to make purchase decisions. These reviews are assumed to be unbiased reflections of other consumers' experiences with the products or services. While extensively assumed, the literature has not tested the existence or non-existence of review manipulation. By using data from Amazon and Barnes & Noble, our study investigates if vendors, publishers, and writers consistently manipulate online consumer reviews. We document the existence of online review manipulation and show that the manipulation strategy of firms seems to be a monotonically decreasing function of the product's true quality or the mean consumer rating of that product. Hence, manipulation decreases the informativeness of online reviews. Furthermore though consumers understand the existence of manipulation, they can only partially correct it based on their expectation of the overall level of manipulation. Hence, vendors are able to change the final outcomes by manipulating online reviewers. In addition, we demonstrate that at the early stages, after an item is released to the Amazon market, both price and reviews serve as quality indicators. Thus, at this stage, a higher price leads to an increase in sales instead of a decrease in sales. At the late stages, price assumes its normal role, meaning a higher price leads to a decrease in sales. Finally, on average, there is a higher level of manipulation on Barnes & Noble than on Amazon.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The rapid adoption of Web 2.0 has unleashed a wave of innovations that might change the way customers acquire information to make product purchases or stock investment decisions. The growth of Web 2.0 has enabled consumers to post reviews describing their experiences with products, product vendors, or service providers and make them available to other prospective consumers. In fact, the marketing literature suggests that consumers depend on online product reviews to make purchase decisions [3,5]. Capital markets research has revealed that the information conveyed by stock message boards are used by investors [1], and a shock to the message board postings is negatively associated with future stock returns [12].

Since consumers increasingly depend on information released through social online channels, such as consumer-generated content, to make product or services purchase decisions, the quality and truthfulness of information available to them is important. Do various entities, such as companies, vendors, publishers, or writers, actively engage in word-of-mouth manipulation, either directly or indirectly, with the goal of changing consumers' final decisions? Such practices are not new for information released through traditional information channels. For example, a rich earnings management literature has revealed that managers deliberately misrepresent financial reports in order to smooth their firm's income, meet a pre-specified target, and get better compensation.

We define review fraud as occurring when online vendors, publishers, or authors write "consumer" reviews by posing as real customers. An email interview with Jonathan Carson, CEO of BuzzMetrics, reveals that promoting new CD releases through chat promotion is almost an industry standard [11]. Such a practice exists even for highly reputable vendors, such as Amazon. In April 2004 James Marcus, a former senior editor for Amazon.com, wrote an alarming article in *The Washington Post* to discuss review fraud. Based on an analysis of reviews of just a few thousand reviewers, he found that a large number of authors on Amazon had got favorable reviews from their friends, relatives, colleagues or paid professionals. In some

cases, these authors even wrote reviews for their own books.[1] Furthermore, such fraud has caused financial loss to society as well.[2]

Recent research concludes that word-of-mouth (WOM) communication is a valuable marketing resource for consumers and marketers with critical implications for a product's success. This literature provides useful insights by linking online reviews with sales. It shows a positive correlation between the average review score and product sales [4–6]. However, there is one implicit but essential assumption in this literature that researchers take for granted as being true, which is:

Assumption 1. Online reviews are written by actual previous customers, not publishers or vendors, etc. Therefore, online reviews reflect either the actual product quality or the product's relative true quality.

If the above assumption is true, then online reviews should reflect a products' true quality; or, all other information (e.g., price, product category, manufacturer, vendor, and shipping terms) being the same, a product with a higher mean consumer product rating should be assumed to have higher quality. This assumption is crucial in justifying the linkage between online reviews and sales. However, the existence of review fraud would invalidate such an assumption and cast doubts on the association between product quality and consumer reviews. If online reviews are indeed written by actual previous customers, then online reviews can help new customers reduce the uncertainties involved in inferring product quality, thus resulting in an increased conversation rate and higher sales. However, if online vendors, publishers, and authors are all able to write "consumer" reviews, then instead of being an uncertainty "reducer", online reviews might become an uncertainty enhancer. In such a case, consumers' beliefs about product quality and vendor reputations derived from online reviews might be totally misleading.

To date, there have been a few analytical studies investigating review fraud [2,11]. Drawing on the observation that the music industry is known to hire professional marketers to write favorable consumer opinions to promote the sales of new albums, Mayzlin [11] built an analytical game theory model in which two competing firms send anonymous messages recommending their own products. Dellarocas [2] analytically shows that if every firm's manipulation strategy monotonically increases with regard to that firm's true quality, then manipulation of online reviews increases the informativeness of online reviews. Under such a circumstance, manipulation increases the separation of the distributions of ratings and will help consumers make better purchase decisions. Even if there is manipulation, consumers are smart and can adjust their interpretation of online opinions accordingly [2]. Combining the implicit assumption stated above (Assumption 1) with these analytical works, we have the following revised assumption based on previous literature:

Assumption 2. Online reviews are written by actual previous customers and not publishers or vendors. Even if there is manipulation, consumers are smart and can adjust their interpretation of online opinions accordingly [2]. Further, as long as the manipulation is monotonically increasing with regard to a product's true quality (i.e., if it is more likely for higher quality vendors to engage in review manipulation), then online reviews with the existence of review fraud are even more informative than when there is no review fraud.

If consumers are indeed smart and if the manipulation is monotonically increasing with respect to (w.r.t) to product quality,

then we need not worry about empirically testing manipulation of online reviews because under such a circumstance, online reviews are more informative. However, are these assumptions true?

In this paper, we analytically and empirically study temporal behaviors of online reviews and address the following research questions:

- Does review fraud actually exist? Is review manipulation a prevalent phenomenon or does it just happen occasionally?
- What types of vendors are more likely to manipulate online reviews: those selling high-quality products or those selling low-quality products? Vendors that receive higher average ratings for their products, or those with lower average ratings?
- Are consumers smart enough to filter out the manipulation as Dellarocas [2] suggests? Are they able to correct for this bias in their purchase decisions? What quality indexes do they use to make purchase decisions in view of the existence of review fraud?
- Is online review fraud a common phenomenon across different websites?

This paper proceeds as follows. Section 2 studies the mean-reverse phenomenon of consumer reviews to motivate our study. By studying the temporal patterns of online reviews, we show that there might be two potential drivers which are consumer taste difference and/or review manipulation that force rating decreases over time. As a nature follow-up question of Section 2, Section 3 answers whether a pure consumer taste difference without manipulation can be the sole underlying driving force. We conclude that we cannot rule out manipulation as one of the potential drivers. The temporal patterns of online reviews can be either driven by pure manipulation or by a joint force of consumer taste difference and manipulation. Section 4 seeks to answer the question of whether low-quality or high-quality vendors are more likely to manipulate consumer reviews. Section 5 analyzes whether consumers correct for manipulation bias when making purchase decisions. Section 6 answers how customers make purchase decisions when manipulation exists. Section 7 checks the robustness of our findings by comparing the online review manipulation between Amazon and Barnes & Nobel. Section 8 contains discussion of the findings, their implications, and some concluding remarks.

2. How do consumer reviews evolve over time?

We study the time-series property of online consumer reviews based on empirical data collected from Amazon.com to reveal why we suspect that vendors, publishers, and authors might consistently manipulate online reviews. Before we discuss our analytical and empirical models, we first discuss where and how we collected our data.

2.1. Data

We collected our data from Amazon Web Service (AWS) and constructed two datasets to examine our research questions. The first dataset is cross-sectional data composed of a random sample of books, DVDs, and videos. For this dataset, we collected the product information and corresponding consumer reviews from Amazon.com in July 2005.[3] The second dataset is a panel dataset composed of a sequence of online review information (price, sales, and review information) for a sample of books, DVDs, and videos collected over several months at approximately three-day intervals. The initial items in this panel dataset were randomly chosen from Amazon in July 2005. For the panel data collection, since it occurs approximately every three days, we identified each data collection batch by a unique sequence number. Because we need to know both the true product quality and the perceived product quality that consumers used to make purchase decisions, we used the

---

[1] http://www.washingtonpost.com/ac2/wp-dyn/A61073-2004Apr8?language= printer.

[2] According to http://www.clickfraudreport.com/1.html, the essence of click fraud is "any click where there is no intention by the clicker to purchase, browse or gain information from the website they visit. And the only goal of a click is to either to drain your marketing budget or generate revenue from the click". Even though we cannot find a dollar amount lose due to review fraud, we believe it is comparable to click fraud.

[3] This study is based on data collected in July 2005. We performed similar data analysis using data collected in February, March, and April of 2005, which rendered similar results.