

Article

# Detection of Fake Reviews: Analysis of Sellers' Manipulation Behavior

Lirong Chen <sup>1,2</sup> , Wenli Li <sup>1,\*</sup> , Hao Chen <sup>3</sup> and Shidao Geng <sup>4</sup>

<sup>1</sup> Faculty of Management and Economics, Dalian University of Technology, Dalian 116024, China

<sup>2</sup> School of Computer Science, Inner Mongolia University, Hohhot 010021, China

<sup>3</sup> School of Business, Qingdao University, Qingdao 266071, China

<sup>4</sup> School of Maritime Economics and Management, Dalian 116026, China

\* Correspondence: wlli@dlut.edu.cn

Received: 2 August 2019; Accepted: 28 August 2019; Published: 3 September 2019



**Abstract:** Online reputation systems play an important role in reducing consumers' purchase uncertainty in online shopping. However, some sellers manipulate reviews for their own interests, which reduces the effectiveness of the reputation system. Unlike the previous studies, which focus on features of reviews and reviewers, this study establishes a game model to analyze sellers' manipulation behavior and identifies what kind of sellers or under what scenario sellers are motivated to manipulate reviews. Our study provides a new perspective for platform to detect fake reviews and helps consumers to make good use of online reviews without getting trapped in some sellers' fraudulent manipulation.

**Keywords:** reputation; game theory; fake-review detection; motivation; online review manipulation

## 1. Introduction

Reputation systems are widely used in various e-commerce platforms, such as Taobao, eBay, and Amazon. Such a system collects user-feedback from consumers and provides reputation information regarding the seller to help potential consumers make their purchase decisions [1–3]. 91% of respondents indicated that they would refer to consumer reviews before shopping online. Positive reviews and high product rating will persuade consumers to buy, while the negative ones will reduce consumers' purchase intention [4]. In addition, negative reviews have a greater influence on consumers' purchase decisions than positive reviews [5].

The value of online consumer reviews offers strong incentives for review manipulation, which results in large numbers of fake favorable reviews [6]. These review manipulations can take many forms. For example, sellers or some interest parties provide online reviews disguised as consumers. Amazon inadvertently leaked the identity of a reviewer who was the author of the book [7]. Sellers harness human laborers to submit fake promotional reviews to improve their stores' reputations [8]. After collecting 2.14 million reviewers and 5.8 million online reviews from Amazon's website, Jindal and Liu [9] found widespread manipulation of reviews. For example, reviews that were submitted by the same commenter for different products are the same or very similar or reviews submitted by different reviewers for the same product or different products are the same or very similar, suggesting that the one person authored many reviews, either as one user or registered as different user names. Obviously, these reviews are inauthentic [10,11].

Consumer reviews are the reputation system's input. The existence of manipulation activity will directly affect the reputation system's effectiveness. According to an investigation, 80% of consumers doubt the authenticity of online reviews [8]. Untruthful reviews considerably bother online consumers. Therefore, the detection of fake online reviews has been a hot topic of theoretical and practical research.

Previous studies have focused on identifying which reviews are fake to detect fake online reviews, trying to identify characteristics of fake reviews or of fake reviewers and then filtering them out. However, the review attackers are also trying to avoid being identified [12]. Researchers found that language in fake reviews were similar to that used in authentic reviews [13]. Therefore, the identification through reviews or reviewers is challenging. Afterwards, we can trace back to the source of these fake reviews. Generally, consumers won't lie for no reason. Accordingly, the real source of fake reviews is the sellers. This study sets aside the review and the reviewer, and starts from the review manipulators—sellers—to establish a game model to analyze the behavior of sellers' review manipulation and identify whether sellers have the probability to manipulate online review. This can provide a new perspective for fake review detection.

## 2. Literature Review

Some sellers try to manipulate online reviews to boost their product sales since online reviews have the power of influencing consumers' purchase decisions. The phenomenon of fake reviews has attracted more and more attention from scholars and practitioners. Researchers are focusing on how to detect fake reviews. In practice, major e-commerce platforms have also adopted some algorithms to filter fake comments. Two factors help to detect the fake comments: the review itself and the reviewer. Most methods for fake-review detection involve machine learning to classify reviews into fake and authentic categories. Researchers have sought to identify attributes to distinguish authentic and fake reviews. They include text-based features—review length, rating, emotion, readability, subjectivity, style of writing, etc. [13,14]—and product-rating characteristics—the difference between a reviewer's rating and the average rating of the product and the rating differences between reviews of different products from the same reviewer [13,15]. Similarly, Deng et al. proposed 11 linguistic deception cues in three categories: word frequency, information richness, and information credibility [16]. However, on the other hand, purveyors of fake reviews work to imitate authentic reviews to avoid being identified [12]. Studies have shown that some fake comments are similar in language to real comments [17]. Therefore, it will be more and more difficult to identify fake reviews through the reviews themselves.

Another idea for identifying fake reviews is based on the reviewers' behavior characteristics: the average number of comments posted by reviewers every day, the time interval between the first and the last comment from a reviewer, the proportion of the first review of products in all the comments posted by a reviewer, the number of votes obtained, and the provision of video information [13,18,19]. Zhang et al. have achieved good results in identifying fake reviews by combining information of review text and reviewer behavior [20].

However, there is a problem in detecting fake reviews using machine learning method. Since review manipulation is covert and invisible, and only the review providers or the operators really know whether a review is fake or not [21]. It is difficult to prepare the training and test data set of fake and authentic reviews. Most researches adopt artificial methods to construct training and test sets [9,22]. For example, Deng et al. asked reviewers to write some true and fake reviews as the training set [16]. Zhang et al. used reviews that were filtered out by the platform as fake comments [20]. The acquisition of appropriate training-set sample data is the main factor influencing the performance of fake review detection.

Fake review detection is similar to crime detection. In addition to collecting clues at the scene of a crime, police also have to infer who has a motive for this crime. Inspired by this, we are going to analyze sellers' review manipulation behavior to help in fake review detection. Some studies are focused on review manipulation behavior. It is found that seller-reputation-escalation (SRE) had been a profitable underground business. Xu, Liu, Wang and Stavrou [8] investigated the impact of this SRE service on reputation escalation and found that this service can boost a seller's reputation at least 10 times faster than legitimate ones. Mayzlin [23] proposed that the cost of review manipulation (mainly the reputation risk) determines the amount of manipulation activity. For example, branded chain hotels

are less likely to engage in review manipulation than independent hotels, because they may suffer great reputation risk if manipulation activity was exposed publicly. Hu, Bose, Gao and Liu [6] developed a review manipulation proxy to identify what kind of sellers may engage in review manipulation based on the discretionary accrual-based earnings management framework.

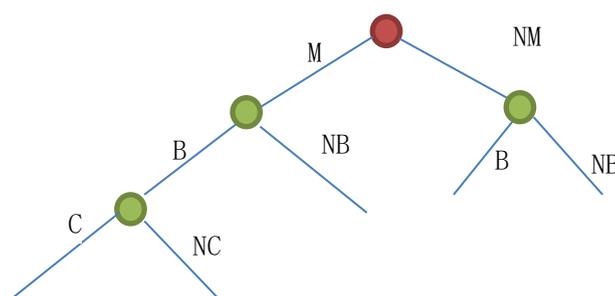
Since this manipulation behavior is hidden, we choose a game model to analyze sellers' motivations. If sellers are motivated to manipulate reviews, fake reviews will be generated. Through the game analysis of review-manipulation behavior, this paper discusses what kinds of sellers or under what scenarios sellers may be so motivated.

This paper proceeds, as follows. In Section 3 we set up game models to analyze sellers' review manipulations. In Section 4, we discuss the implications for management. In Section 5, we conclude the paper and discuss some limitations and future research.

### 3. Game Model

In this study, sellers' review manipulation refers to the manufacturing of fake favorable reviews in various ways to improve sellers' own reputation, excluding malicious negative comments on competing sellers.

The game model that was established in this paper includes sellers, platforms, and consumers. Sellers manipulate online reviews to improve the reputation score and attract consumers to buy, thus gaining extra profit. Consumers form their expectations according to sellers' reputation scores and product price. If consumers find that the goods are not consistent with their expectations, they will believe that sellers cheat consumers with false advertising, and they may complain to the platform. If complaints are verified by the platform, sellers will be fined or prohibited from continuing business on the platform. Some platforms have their own filtering algorithms for fake transactions or spam reviews. Yelp, for example, has a filtering algorithm that removes reviews that are not related to the product or present information, like advertisements. Some platforms, such as JD.com, do not filter reviews. Moreover, platforms have great difficulty in identifying manipulated information. To simplify the game process, we assume that the platform takes no initiative to investigate sellers' manipulation behavior, but punishes sellers based on consumer complaints. Therefore, the main players in the game are simplified as sellers and consumers. It is unknown whether sellers have provided false information during product introductions or manipulated reviews in advance. The game proceeds, as shown in Figure 1.



**Figure 1.** Game of review manipulation or not by sellers (platform does not filter reviews in advance).

The red decision node represents the sellers' decision choices. Sellers have two strategies: manipulating reviews (M) or not (NM). The green decision nodes indicate consumers' decisions. Consumers have two strategies, buy (B) or not (NB), and then post-purchase complaint (C) or not (NC). The platform will punish sellers if complaints are verified.

#### 3.1. Consumer's Choice

Whether consumers choose to buy depends on the consumers' expected net utility. One aspect is the satisfaction brought by the expected quality of goods; the other aspect is the price paid [24,25].

Suppose that the expected net utility of a consumer is  $U$ , then  $U = \theta q^e - p$ , where  $\theta$  is the consumer's preference for product quality, which is uniformly distributed,  $0 \leq \theta \leq 1$ .  $p$  is the price of the product,  $q^e$  is consumer's expected product quality.

After receiving the products, consumers will submit an overall rating of the product according to their experiences. The product's rating is based on the comparison between the actual utility and expected utility of the goods received by consumers. Therefore, the rating reflects both the quality and the price information [26]. Subsequently, potential consumers' expected quality for the goods is related to both the rating and the price. Higher rating leads to higher expected quality; higher prices lead to higher expected quality. To simplify the calculation, we assume there is a linear relationship between the expected quality, product ratings (usually average ratings in the system), and price:  $q^e = \beta_0 R + \beta_1 p$ ,  $\beta_0 > 0$ ,  $\beta_1 > 0$ .  $R$  is the average product rating (reputation score) in the system. Subsequently,

$$U = \theta q^e - p = \theta(\beta_0 R + \beta_1 p) - p$$

Only when  $U > 0$ , consumers will choose to buy. If a consumer finds that the actual quality of product is not consistent with seller's description or deviated a lot from his expected quality, he will choose to complain to the platform, and the seller will be punished by platform.

### 3.2. Seller's Choice

Seller's manipulation of online review depends on the benefit and cost of this activity. As online reviews become an important form of word-of-mouth communication, they become increasingly important to consumers' purchasing decisions. Consumers are more likely to trust reviews from strange consumers than from commercial advertisements. Numerous studies in different fields have shown that online reviews have a significant impact on business sales [2,27–32]. As sales increase, so will profits. This is the main reason that sellers manipulate fake reviews to improve their reputation [33,34]. Higher profits from review manipulation will encourage more manipulation activities by the sellers.

The cost of review manipulation mainly includes two parts. First, the direct cost of employing someone to write favorable reviews. For example, sellers may hire ghostwriters to create favorable reviews, which is equivalent to buying favorable reviews. Second, the cost of being punished by the platform. Consumers may complain to the platform if the product is found to be inconsistent with their expectation or deviated from the description by the sellers. After verification, the platform will punish the sellers. Accordingly, sellers who manipulate reviews are faced with the risk of consumer complaints and punishment by the platform. This risk is related to the deviation between the actual reputation score and score after the sellers' manipulation. If sellers manufacture fake comments and make the score deviate from the actual high, the risk of consumer complaints will increase. The higher the punishment degree is, the higher the cost of review manipulation. For example, some platforms will prevent sellers with misleading propaganda from trading on the platform. If sellers choose to continue selling on the platform, they have to register an account again, and the accumulation of reputation needs to start from scratch. For sellers with a certain basis of reputation, the manipulation cost is relatively high.

Accordingly, for sellers, on the one hand, a high reputation score will attract more consumers and increase online sales. On the other hand, sellers must pay for a reputation-score increase. Therefore, sellers must make a tradeoff between gaining a larger market share by manipulating reviews and the possible costs of doing so.

Some goods are durable goods, and there are hundreds of sellers that provide similar goods on e-commerce platforms. Some sellers face one-off purchases. Sellers with nondurable goods or few potential consumers rely more on repeat purchases. Accordingly, we divide the game between sellers and consumers into a one-off game and a repeated-purchase game.

### 3.3. A One-Time Game Model of Seller's Review Manipulation

According to the degree of product substitutability, the discussion is divided into two cases, weak competition among sellers, wherein each seller has its own relatively independent demand market and fierce competition among sellers, such as different sellers selling products of the same brand and facing the same market.

Here is the reason why we consider the scenario that seller facing independent market demands. Only when consumer's expected utility is above zero will the consumer choose to buy. Even though there is no competitor, the sellers hope to improve consumers' expected utility and attract more consumers. We can also view this situation as a prerequisite for sellers to compete with others. If a seller cannot meet minimum requirements of the consumer, he has no chance to compete with other sellers.

#### 3.3.1. Sellers Facing Independent Market Demands

Suppose that consumers have unit demand for a product. That is, they consume up to one unit of the product, regardless of its quality. The initial true reputation score of the seller's product is  $r_1$ , and the initial price is  $p_1$ . We suppose that the marginal cost of product is zero for simplicity. Consumers usually infer the quality of the product according to the reputation score and price of the product. Sellers do not change the price. If a seller does not manipulate reviews, the consumer's expected utility is

$$U = \theta(\beta_0 r_1 + \beta_1 p_1) - p_1,$$

Consumers whose utility is greater than zero choose to buy, and the demand for goods is  $1 - \frac{p_1}{\beta_0 r_1 + \beta_1 p_1}$ . The purpose of sellers' review manipulation is to rapidly improve the reputation. For example, through the accumulation of fake ratings, the reputation score of the product rises to  $r_2$ , the target reputation. Whether consumers complain is related to the degree of the seller's review manipulation. We can measure the degree of manipulation by the difference between the initial reputation and the target reputation score. The more the reputation score increased, the greater the possibility of the consumers complaining. Assume that the profit when sellers do not manipulate the review is  $\pi_1$  and the expected profit after manipulation is  $\pi_2$ ,

$$\pi_1 = p_1 \left( 1 - \frac{p_1}{\beta_0 r_1 + \beta_1 p_1} \right),$$

$$\pi_2 = p_1 \left( 1 - \frac{p_1}{\beta_0 r_2 + \beta_1 p_1} \right) - \alpha n (r_2 - r_1) - T \cdot (r_2 - r_1).$$

In the above formula,  $T$  is the penalty imposed by the platform when the platform receives consumer complaints of the seller,  $\alpha$  is the review-manipulation cost coefficient, and  $n$  is the product's current number of reviews. The cost of manipulating is related to the degree of reputation increase and the number of reviews already received. The greater the reputation increased, the higher the cost of manipulation; the more reviews already received, the higher the unit cost of manipulation. For example, a seller has gained 10 reviews until now, and the reputation score is 4, if this seller wants to increase the reputation score to 4.5, he has to manipulate 10 fake reviews with each review being rated 5. However, if this seller has gained 20 reviews, he needs 20 fake reviews to achieve an average score of 4.5. With the increase of  $n$ , the unit cost of manipulation will increase too. The game strategy is described in Table 1.

**Table 1.** Game model of sellers’ review manipulation behavior.

		Seller	
		No Manipulation	Manipulation
Consumer	Buy	$\theta q_1^e - p_1, \pi_1$	$\theta q_2^e - p_1, \pi_2$
	Not buy	0, 0	$0, -\alpha n(r_2 - r_1)$

As long as  $\theta q_1^e - p_1 > 0$ , Consumers will choose to buy, and consumers with quality preference more than  $\frac{p_1}{\beta_0 r_1 + \beta_1 p_1}$  will buy the product. When consumers choose to buy, and  $\pi_2 > \pi_1$ , sellers will choose to manipulate online reviews:

$$\alpha \leq \frac{p_1 * p_1}{n(r_1 r_2 + (r_1 + r_2)p_1 + p_1^2)} - \frac{T}{n}.$$

Assume that  $Y = \frac{p_1 * p_1}{n(r_1 r_2 + (r_1 + r_2)p_1 + p_1^2)} - \frac{T}{n}$ . A larger value of  $Y$  leads to a higher upper limit of  $\alpha$  and a greater probability of sellers manipulating reviews. As can be seen from the above inequality, a smaller initial reputation score,  $r_1$ , yields a larger value of  $Y$ , which allows  $\alpha$  to take a larger value, which allows the seller to manipulate online reviews at a higher unit cost. Other things being equal, a large  $r_2$  will result in a small value of  $Y$ . That is, with the increase of the target reputation, the upper limit value of  $\alpha$  is small, which allows a small unit cost of manipulation and reduces the possibility of sellers manipulating online review. In addition, a large  $p_1$  will yield a big  $Y$ . That is, when the product price is higher, the possibility of sellers manipulating will also increase. Similarly, a larger  $n$  yields a smaller  $Y$ : More reviews already being obtained by the product reduces the possibility of review manipulation.

**Corollary 1.** Sellers are more likely to manipulate reviews when the product’s reputation score is low.

**Corollary 2.** If the product’s current reputation score is very high, the seller is less likely to have obtained it through manipulation.

**Corollary 3.** Sellers with high-priced products are more likely to manipulate reviews.

**Corollary 4.** A more severe punishment imposed by the platform on sellers reduces the likelihood they will manipulate reviews.

**Corollary 5.** Products with fewer reviews are more likely manipulated by the seller.

From the above conclusions, we find that Corollaries 1 and 3 are consistent with the research conclusions of Hu et al. [6].

### 3.3.2. Sellers Facing a Competitive Market

Suppose there are two sellers on the platform, A and B. The two sellers operate in a common market and compete for market share. Assume the initial product rating (real rating)  $R_{A1} < R_{B1}$ , the price of product A is less than that of B,  $P_A < P_B$  (to differentiate between the product rating in two scenarios, we use uppercase  $R$  and  $P$  to represent product rating and product price here). Additionally,  $q_A^e = \beta_0 R_{A1} + \beta_1 P_A$  and  $q_B^e = \beta_0 R_{B1} + \beta_1 P_B$ , so  $q_A^e < q_B^e$ ; consumers expect that the quality of product A is less than the quality of product B. The price of product A is less than the price of product B, so each product has some market demand. Consumers with quality preference greater than  $\tilde{\theta} = \frac{P_B - P_A}{\beta_0(R_{B1} - R_{A1}) + \beta_1(P_B - P_A)}$  choose to buy the high-quality product ( $\theta q_B^e - P_B > \theta q_A^e - P_A$ ). Consumers with quality preference less than  $\tilde{\theta}$  and greater than  $\frac{P_A}{\beta_0 R_{A1} + \beta_1 P_A}$  ( $\theta q_A^e - P_A > 0$ ) will buy the

low-quality product. Hence, demands for the high-quality and low-quality products are  $D_A$  and  $D_B$ , respectively, see Figure 1.

$$D_B = 1 - \frac{P_B - P_A}{\beta_0(R_{B1} - R_{A1}) + \beta_1(P_B - P_A)},$$

$$D_A = \frac{P_B - P_A}{\beta_0(R_{B1} - R_{A1}) + \beta_1(P_B - P_A)} - \frac{P_A}{\beta_0 R_{A1} + \beta_1 P_A}.$$

If the low-quality sellers do not manipulate the comments, and the high-quality sellers choose to manipulate the comments and increase the reputation scores, the high-quality sellers will gain more market share. If the high-quality sellers choose not to manipulate and the low-quality sellers manipulate the reviews, the low-quality sellers will gain more market share. It means manipulation by low-quality sellers will make high-quality sellers consider manipulating. Likewise, manipulation by high-quality sellers will affect low-quality sellers' actions. The two sides reach an equilibrium, such that, for example, the high-quality product's reputation score is  $R_{B2}$  and the low-quality product's score is  $R_{A2}$ . The profits of the two types of sellers will be  $\pi_{A2}$  and  $\pi_{B2}$ .

$$\pi_{A2} = \left( \frac{P_B - P_A}{\beta_0(R_{B2} - R_{A2}) + \beta_1(P_B - P_A)} - \frac{P_A}{\beta_0 R_{A2} + \beta_1 P_A} \right) \cdot P_A - \alpha n_A (R_{A2} - R_{A1}) - T(R_{A2} - R_{A1})$$

$$\pi_{B2} = \left( 1 - \frac{P_B - P_A}{\beta_0(R_{B2} - R_{A2}) + \beta_1(P_B - P_A)} \right) \cdot P_B - \alpha n_B (R_{B2} - R_{B1}) - T(R_{B2} - R_{B1})$$

Since we are not concerned with the effects of  $\beta_0$  or  $\beta_1$  on sellers' manipulation behavior, we set  $\beta_0 = \beta_1 = 1$  for simplicity, while simplifying the sellers' profit.

$$\pi_{A2} = \left( \frac{P_B - P_A}{R_{B2} - R_{A2}} - \frac{P_A}{R_{A2}} \right) \cdot P_A - \alpha n_A (R_{A2} - R_{A1}) - T(R_{A2} - R_{A1})$$

$$\pi_{B2} = \left( 1 - \frac{P_B - P_A}{R_{B2} - R_{A2}} \right) \cdot P_B - \alpha n_B (R_{B2} - R_{B1}) - T(R_{B2} - R_{B1})$$

To maximize profits, we obtain the partial derivatives of  $R_{A2}$  and  $R_{B2}$ .

$$R_{A2} = P_A \sqrt{\frac{P_B}{\alpha(n_A P_B - n_B P_A) + T(P_B - P_A)}}$$

$$R_{B2} = P_A \sqrt{\frac{P_B}{\alpha(n_A P_B - n_B P_A) + T(P_B - P_A)}} + \sqrt{\frac{P_B(P_B - P_A)}{\alpha n_B + T}}$$

As long as the reputation score that maximizes the profit of the seller is greater than the actual score at present, the seller is likely to manipulate reviews. That is, if  $R_{A2} > R_{A1}$ , sellers selling product A are motivated to manipulate the reviews. Similarly, if  $R_{B2} > R_{B1}$ , the sellers B are also motivated to manipulate. To simplify the calculation, we assume that seller A and B have the same number of initial reviews,  $n_A = n_B = n$ . For low-quality sellers,  $\alpha < \frac{P_A^2 * P_B}{R_{A1}^2 * n(P_B - P_A)} - \frac{T}{n}$ . When the punishment is heavy, the possibility of manipulation by the sellers is small. In addition, when the initial cost performance of the low-quality product is low, as in Figure 2, the left boundary point of the low-quality product's market demand is further to the right and the market share is low, so sellers are more likely to manipulate reviews. A greater price difference between the low-quality and high-quality products makes sellers less likely to manipulate the reviews. For high quality sellers,  $\alpha < \frac{P_B^3}{R_{B1}^2 (P_B - P_A)n} - \frac{T}{n}$ , a higher price difference lowers the possibility of manipulation. In addition, a higher initial reputation score for high-quality sellers does the same.

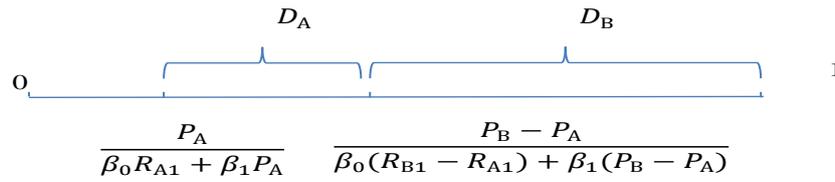


Figure 2. Sellers’ market share in a competitive market.

**Corollary 6.** A greater the price difference between high- and low-quality products makes sellers less likely to manipulate the reviews.

**Corollary 7.** A higher initial reputation score makes seller less likely to manipulate.

**Corollary 8.** More-severe punishment imposed by the platform on sellers deters review manipulation.

Corollary 7 is the same as Corollary 1, and Corollary 8 is the same as Corollary 4.

If  $R_{A1} > R_{B1}$  and  $P_A < P_B$ , for  $q_A^e = \beta_0 R_{A1} + \beta_1 p_A$ ,  $q_B^e = \beta_0 R_{B1} + \beta_1 p_B$ , then the relative expected quality of products A and B is uncertain. When  $\frac{\beta_0}{\beta_1} < \frac{p_B - p_A}{R_{A1} - R_{B1}}$ ,  $q_A^e < q_B^e$  and the game between the two sellers is the same as above. In contrast, if  $\frac{\beta_0}{\beta_1} > \frac{p_B - p_A}{R_{A1} - R_{B1}}$ , then  $q_A^e > q_B^e$ . Since the price of high-quality sellers is lower ( $P_A < P_B$ ), consumers will buy all of product A, and the market demand for product B is zero. To obtain market demand, seller B must improve the reputation score and make  $q_A^e < q_B^e$ . Meanwhile, the market share for seller A will decline, so both of the sellers will choose to manipulate. For the unit manipulation cost for seller A is lower than that of seller B, the latter will not manipulate. Therefore, even if sellers selling low-quality goods intend to sell seconds at best-quality prices, they will not set prices higher than those of the high-quality sellers; they will set slightly lower prices.

**Corollary 9.** If the sellers with low quality products intend to sell seconds at a higher price, they will choose to set slightly lower prices than those of the best-quality sellers.

### 3.4. The Repeated-Game Model of Sellers’ Manipulation Behavior

Some sellers rely on consumers’ repeat purchases due to product characteristics or their operation mode. For example, sellers providing local life services have few potential customers and high sunk costs.

As in the previous analysis, we discuss two cases: two sellers with relatively independent markets and two sellers facing a common market.

#### 3.4.1. Sellers Facing Independent Markets

If the sellers choose not to manipulate the reviews, the quality of the products received by the consumers is consistent with expectations, consumers will not choose to complain, and the profits of the sellers in each period will remain unchanged:  $\pi'_1 = N\pi_1 = Np_1 \left(1 - \frac{p_1}{\beta_0 r_1 + \beta_1 p_1}\right)$ , where  $N$  is the number of transactions between buyers and sellers, as well as the number of games between buyers and sellers.  $\pi'_1$  is the seller’s expected earning when not manipulating.  $p_1$  and  $r_1$  are the initial price and reputation score. If the seller chooses to manipulate, the product’s reputation may improve and attract new consumers to increase sales. On the other hand, the seller must bear the direct cost of review manipulation and it may be the cost of platform punishment because of consumer complaints. If consumers find that the received products are inconsistent with their expectations, they may abandon subsequent transactions with the seller, and the seller will lose these consumers. In addition, as the number of real reviews increases, the seller’s reputation score will return to its real value. The unit cost of manipulation will increase if the seller chooses to improve the reputation score

through manipulation in a subsequent stage. If the seller only chooses to manipulate the comments in the first stage of the game, the market demand will be the same as the market in a one-time game. Therefore, this problem comes down to whether sellers choose to manipulate in the one-time game, so the conclusion is consistent with the conclusions above. If the seller chooses to manipulate reviews at each stage to increase the reputation score from  $r_1$  to  $r_2$ , then the expected profit is

$$\pi'_2 = p_1 \left( 1 - \frac{p_1}{\beta_0 r_2 + \beta_1 p_1} \right) \left( 1 + (1 - \beta) + (1 - \beta)^2 + \dots + (1 - \beta)^N \right) - \sum_{t=1}^N \alpha_t (r_2 - r_1) - N\beta T.$$

In the above formula,  $\alpha_t$  is the unit cost of manipulation in the  $t$ th game. With increasing  $t$ , the reviews accumulate more and more, so  $\alpha_t$  increases with  $t$ .  $\beta$  is the proportion of consumers who find that the products do not conform to expectations and complain in each stage of the game. As  $N$  increases, the expected profit decreases. Therefore, the seller will not choose to manipulate reviews if the seller relies on the repeat customers.

### 3.4.2. Sellers Compete for Market Share

Suppose that two sellers A and B face a common market, the initial reputation score  $R_{A1} < R_{B1}$ , and  $P_A < P_B$ , then  $q_A^e < q_B^e$ . That is, consumers expect the quality of product A is lower than that of product B. If the low-quality sellers manipulate online reviews, while the high-quality sellers choose not to manipulate, the high-quality sellers will lose market share, but do not pay the cost of the manipulation and they are not punished by the platform. If the reputation score of low-quality products rises to  $R_{A2}$ , the expected profit of high-quality sellers in the next  $N$  periods will be

$$\pi_{NB} = \left( 1 - \frac{P_B - P_A}{\beta_0 (R_{B1} - R_{A2}) + \beta_1 (P_B - P_A)} \right) P_B \cdot N.$$

If high-quality sellers choose to manipulate, the consumers find that the received product is inconsistent with expectation, and they may complain. In that case, sellers will be punished by the platform, and they will lose these consumers, assume a ratio,  $\beta$ , of consumers choose to complain and that sellers will be punished. Subsequently, the sellers will lose a proportion of demand in each phase, and the sellers' expected profits in  $N$  periods are

$$\pi'_{NB} = \left( 1 - \frac{P_B - P_A}{\beta_0 (R_{B2} - R_{A2}) + \beta_1 (P_B - P_A)} \right) \left( 1 + 1 - \beta + (1 - \beta)^2 + \dots + (1 - \beta)^N \right) P_2 - \sum_{t=1}^N \alpha_t (R_{B2} - R_{A2}) - N\beta T.$$

After simplification,

$$\pi'_{NB} = \left( 1 - \frac{P_B - P_A}{\beta_0 (R_{B2} - R_{A2}) + \beta_1 (P_B - P_A)} \right) \frac{1}{(1 - \beta)} P_2 - \sum_{t=1}^N \alpha_t (R_{B2} - R_{A2}) - N\beta T.$$

With increasing  $N$ ,  $\pi'_{NB}$  will be smaller and smaller. If high-quality sellers choose not to manipulate, although they may lose part of market share to the other product, they save the cost of manipulation, being punished. Therefore, high-quality sellers relying on repeat customers will choose not to manipulate reviews from the perspective of continuous operation if low-quality sellers choose to manipulate reviews.

Similarly, if the low-quality sellers choose not to manipulate the comments and the high-quality sellers do not manipulate the comments, then the expected profit of the high-quality sellers in the future  $N$  periods is  $\left( 1 - \frac{P_B - P_A}{R_{B1} - R_{A1}} \right) \cdot P_B \cdot N$ . If high-quality sellers manipulate, the expected profit is  $\left( 1 - \frac{P_B - P_A}{R_{B2} - R_{A1}} \right) \cdot \frac{1}{(1 - \beta)} P_B - \sum_{t=1}^N \alpha_t (R_{B2} - R_{A1}) - N\beta T$ . For the aforementioned reasons, high-quality sellers

will choose not to manipulate. Therefore, regardless of whether low-quality sellers choose to manipulate or not, high-quality sellers will choose not to manipulate reviews.

In the case that high-quality sellers choose not to manipulate reviews, the analysis on whether low-quality sellers are motivated to manipulate reviews is similar to the analysis on that of high-quality sellers that is mentioned above. The final equilibrium result is that both low- and high-quality sellers choose not to manipulate reviews.

**Corollary 10.** *The sellers have less incentive to manipulate reviews when they rely on repeat customers.*

#### 4. Management Implications

Generally, in a C2C online market, sellers are facing one-time transaction with the same consumer. An increase of reputation or product rating will attract more consumers and more profits. Accordingly, sellers are motivated to increase their reputation score, even by review manipulation. However, the cost of manipulation and risk of being exposed hinder this activity. When the product rating is very low, the benefit from manipulation is higher, and the sellers are more likely to provide fake favorable reviews. However, these sellers will not increase the product rating too much, otherwise, consumers can easily find that the products they receive are not in accordance with their expectations or the description of the sellers. Then sellers may be punished by platform.

In our model, we assumed that the marginal cost of product is zero for simplicity and found that sellers with high-priced product are more likely to fake. If we let go this assumption, it means that if net profit of product is higher, sellers are more likely to buy positive reviews. It is a kind of advertising investment. When there are few reviews for a product, the cost of increase product rating is lower, sellers are probably to fake favorable reviews. For example, when a new seller comes, or a new product just hit on the shelves, it has few reviews. Subsequently, manipulation at this time cost fewer, and sellers are more likely to manufacture positive reviews for themselves. Similarly, a heavy punishment will prevent sellers from participating in such kind of activities.

Things get more complicated when competitors are considered. If a competitor chooses to manipulate the online reviews, the seller is then forced to produce fake reviews; otherwise, his market share will be squeezed. Especially when product prices are close to each other, sellers may fall in a race of review manipulation. Through an equilibrium analysis, we found that although a low-quality seller wants to pretend to be high quality seller, generally, he will not set a higher price than high quality sellers.

If a seller relies on repeat customers, once a consumer feels cheated, he will not buy again, the seller loses this consumer forever. Accordingly, there is no need for this kind of sellers to attract customers with fake reviews. According to our model, a high-quality new seller can draw customers' attention through significant low price or even free trial. This is just the familiar O2O (online to offline) model. For local service providers, usually they have limited potential consumers, and most of them have brick-and-mortar stores, and it means there is high sunk cost. Hence, sellers need to keep customers in the long-term operation to make a profit. Afterwards, we can infer that those local service sellers on Meituan.com (O2O model) are less motivated to manipulate the online reviews.

Our paper presents some corollaries that give very important implications for different participants (buyers, sellers, and the third-party e-commerce platform) in e-commerce transaction.

Our study provides a new idea for the detection of fake reviews. It is difficult for consumers to detect the fake reviews through ratings and review content, as all of the fake reviews are written either using an assumed customer name or an anonymous identity. However, we can speculate on sellers' motive to manipulate the reviews. If a seller has the motive to manipulate the reviews, the reputation score of the product may be biased and misleading. Subsequently, consumers should form a new quality assessment accordingly [11]. By analyzing sellers' motivations for manipulation, consumers

can gain an understanding of when manipulation is more likely to happen and adjust their perception of products' quality to make a wise purchase decision.

Trust is a key factor for successful transactions between buyers and sellers in the e-commerce environment. Sociologists and economists classify trust into three categories: trust that is based on personality traits, trust based on reputation, and trust based on institutions [35]. Through the game analysis of sellers' manipulation of online reviews, we have found that, in the e-commerce environment, it is difficult for consumers to trust some sellers only through reputation that is gathered exclusively from online reviews. For example, in the C2C trading mode, it is difficult to gain full trust from consumers to promote the sales of high-priced goods only by reputation scores. When such a reputation is not enough to establish trust from buyers, it may be necessary to supplement and form consumer trust by other means. For example, seven-day no-reason returns, free return insurance provided by sellers, and the flag of an official authorized store.

This paper also has implications for reputation system design. Our results suggest that products with fewer reviews are more likely manipulated by the seller. On the one hand, fewer favorable online reviews will not convince consumers that the product is of high quality, so it is necessary to incentivize consumers to provide more authentic reviews. More truthful reviews will gradually reduce the impact of fake reviews on reputation scores [11]. On the other hand, for the product that has gained a number of reviews already, the latest reviews should be the most valuable, because they are more likely from the true consumers. Therefore, online reviews should be chronological in the system. Although consumers can choose to sort the display of reviews by time in the system, reviews are sorted by recommendation defaulted.

In addition, the platform should take a heavier punishment for the sellers who manipulate online reviews, especially those that provide shoddy products. In fact, there are two kinds of fake reviews, one is promotional reviews from sellers with high quality products. High quality sellers are just advertising to let more consumers know about their products, and the reputation score is consistent with product's true quality. Subsequently, these sellers would not be punished in general. The other is manipulated reviews from sellers with low quality products. These reviews may mislead consumers' purchase decisions, and sellers should be heavily punished.

## 5. Conclusions and Future Research

The existence of fake reviews may affect the validity of a reputation system and mislead consumers' purchase decision making. However, due to the characteristics of fake reviews, it is difficult to identify them. We find that sellers with the following characteristics are less motivated to manipulate online reviews, based on our game analysis of sellers' review manipulation behaviors in different situations: sellers who rely on repeat purchase, sellers facing a stiff penalty for manipulation, sellers selling lower-priced products when competing products are much more expensive, sellers whose reputation score is already high. Through the analysis of the sellers' manipulation behavior, on the one hand, consumers can avoid blindly depending on reputation information to make purchase decisions and, on the other hand, the sellers should clarify the role the reputation mechanism plays in product marketing.

There are some limitations in our work. First, in our model, we assume product rating and price influence consumers' expectation of quality and then their purchase decision, thus the benefit from manipulating. However, in reality, there are other factors that affect consumers' expected quality. For example, a long time operating on the platform is also a signal of product quality. Accordingly, the model can be further improved considering other factors. Similarly, online review and product rating may play different roles in purchasing search goods and experience goods, so sellers with different kinds of product may differ in their manipulation activity. Moreover, this may be an interesting issue for future work.

Second, we only infer sellers' manipulation activity through an analytical model. In our paper, we give an explanation that we cannot directly observe the manipulation due to the nature of this behavior. However, if the data about those sellers that are complained by consumers can be collected

or acquired, future research can examine our model through an empirical study and it may come up with other manipulation indicators.

**Author Contributions:** Conceptualization, L.C. and W.L.; supervision of the research, W.L.; writing, revision, and finalization of the manuscript, L.C.; criticism and revision of manuscript, H.C. and S.G.

**Funding:** This paper was funded by the National Natural Science Foundation of China (71862027,71874022 and 71431002).

**Acknowledgments:** The authors are grateful to Paulo Goes, Alok Gupta, Elena Karahanna, Arun Rai for their valuable help and guidance on MIS author development workshop. It has dramatically improved the quality of this paper.

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

## References

1. Chen, Y.; Xie, J. Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Manag. Sci.* **2008**, *54*, 477–491. [CrossRef]
2. Öğüt, H.; Onur Taş, B.K. The influence of internet customer reviews on the online sales and prices in hotel industry. *Serv. Ind. J.* **2012**, *32*, 197–214. [CrossRef]
3. Floyd, K.; Freling, R.; Alhoqail, S.; Cho, H.Y.; Freling, T. How online product reviews affect retail sales: A meta-analysis. *J. Retail.* **2014**, *90*, 217–232. [CrossRef]
4. Anderson, E.W. Customer Satisfaction and Word of Mouth. *J. Serv. Res.* **1998**, *1*, 5–17. [CrossRef]
5. Nam, S.; Manchanda, P.; Chintagunta, P.K. The Effect of Signal Quality and Contiguous Word of Mouth on Customer Acquisition for a Video-on-Demand Service. *Mark. Sci.* **2010**, *29*, 690–700. [CrossRef]
6. Hu, N.; Bose, I.; Gao, Y.; Liu, L. Manipulation in digital word-of-mouth: A reality check for book reviews. *Decis. Support Syst.* **2011**, *50*, 627–635. [CrossRef]
7. Smith, D.A. Amazon Reviewers Brought to Book. Available online: <https://www.theguardian.com/technology/2004/feb/15/books.booksnews> (accessed on 26 August 2019).
8. Xu, H.; Liu, D.; Wang, H.; Stavrou, A. E-commerce reputation manipulation: The emergence of reputation-escalation-as-a-service. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; pp. 1296–1306.
9. Jindal, N.; Liu, B. Review spam detection. In Proceedings of the 16th International Conference on World Wide Web, Banff, AB, Canada, 8–12 May 2007; pp. 1189–1190.
10. Wang, Y.; Lu, X.; Tan, Y. Impact of product attributes on customer satisfaction: An analysis of online reviews for washing machines. *Electron. Commer. Res. Appl.* **2018**, *29*, 1–11. [CrossRef]
11. Chen, L.; Jiang, T.; Li, W.; Geng, S.; Hussain, S. Who should pay for online reviews? Design of an online user feedback mechanism. *Electron. Commer. Res. Appl.* **2017**, *23*, 38–44. [CrossRef]
12. Lappas, T. Fake reviews: The malicious perspective. In Proceedings of the International Conference on Applications of Natural Language Processing and Information Systems, Groningen, The Netherlands, 26–28 June 2012; pp. 23–34.
13. Mukherjee, A.; Venkataraman, V.; Liu, B.; Glance, N. What yelp fake review filter might be doing? In Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, Cambridge, MA, USA, 8–11 July 2013.
14. Banerjee, S.; Chua, A.Y. A study of manipulative and authentic negative reviews. In Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication, Siem Reap, Cambodia, 9–11 January 2014; p. 76.
15. Xu, C. Detecting collusive spammers in online review communities. In Proceedings of the Sixth Workshop on Ph. D. Students in Information and Knowledge Management, San Francisco, CA, USA, 27 October–1 November 2013; pp. 33–40.
16. Shasha, D.; Pengzhu, Z.; Xiaoyan, Z.; Xinmiao, L. Deception Detection based on Fake Linguistic Cues. *J. Syst. Manag.* **2014**, *23*, 263–270.
17. Chen, Y.-R.; Chen, H.-H. Opinion Spam Detection in Web Forum: A Real Case Study. In Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, Florence, Italy, 18–22 May 2015; pp. 173–183.

18. Kamberer, D. Understanding the Yelp review filter: An exploratory study. *First Monday* **2014**, *19*. [[CrossRef](#)]
19. Mukherjee, A.; Kumar, A.; Liu, B.; Wang, J.; Hsu, M.; Castellanos, M.; Ghosh, R. Spotting opinion spammers using behavioral footprints. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 632–640.
20. Zhang, D.; Zhou, L.; Kehoe, J.L.; Kilic, I.Y. What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. *J. Manag. Infor. Syst.* **2016**, *33*, 456–481. [[CrossRef](#)]
21. Luca, M.; Zervas, G. Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. *Manag. Sci.* **2016**, *62*, 3412–3427. [[CrossRef](#)]
22. Ott, M.; Choi, Y.; Cardie, C.; Hancock, J.T. Finding deceptive opinion spam by any stretch of the imagination. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; Volume 1, pp. 309–319.
23. Mayzlin, D. Promotional Chat on the Internet. *Mark. Sci.* **2006**, *25*, 155–163. [[CrossRef](#)]
24. Dellarocas, C. Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Manag. Sci.* **2006**, *52*, 1577–1593. [[CrossRef](#)]
25. Tirole, J. *The Theory of Industrial Organization*; MIT Press: Cambridge, MA, USA, 1988.
26. Xinxin, L.; Hitt, L.M. Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis. *MIS Q.* **2010**, *34*, 809–831.
27. Sun, M. How does the variance of product ratings matter? *Manag. Sci.* **2012**, *58*, 696–707. [[CrossRef](#)]
28. Melnik, M.I.; Alm, J. Seller Reputation, Information Signals, and Prices for Heterogeneous Coins on eBay. *South. Econ. J.* **2005**, *72*, 305–328. [[CrossRef](#)]
29. Liu, Y. Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *J. Mark.* **2006**, *70*, 74–89. [[CrossRef](#)]
30. Li, X.; Hitt, L.M. Self-selection and information role of online product reviews. *Inf. Syst. Res.* **2008**, *19*, 456–474. [[CrossRef](#)]
31. Gu, B.; Park, J.; Konana, P. Research note—the impact of external word-of-mouth sources on retailer sales of high-involvement products. *Inf. Syst. Res.* **2012**, *23*, 182–196. [[CrossRef](#)]
32. Chen, H.; Duan, W.; Zhou, W. The interplay between free sampling and word of mouth in the online software market. *Decis. Support Syst.* **2017**, *95*, 82–90. [[CrossRef](#)]
33. Xiong, J.-Y.; Zhong, Y.S. Fraud against model for seller’s reputation of C2C. *Comput. Sci.* **2012**, *39*, 68–71.
34. Hendrikx, F.; Bubendorfer, K.; Chard, R. Reputation systems: A survey and taxonomy. *J. Parallel Distrib. Comput.* **2015**, *75*, 184–197. [[CrossRef](#)]
35. Yang, J.; Zhang, W.; Zhou, L. The Complementarity and Substitution of Reputation and Regulation: An Empirical Study Based on Online Transaction Data. *Manag. World* **2008**, *07*, 18–26. [[CrossRef](#)]

