How Should We Deal with Malicious Customers' Threats in Online Review? Perspectives of Retailers, Customers, and Platforms

Hemant K. Bhargava
University of California Davis
hemantb@ucdavis.edu

Junghee Lee
University of Notre Dame
jlee93@nd.edu

Xinxue Qu
University of Notre Dame
xqu2@nd.edu

Daewon Sun
University of Notre Dame
dsun@nd.edu

Abstract

We show that existence of malicious customers will distort the retailer's overall rating and under certain conditions, complying to malicious customers' request can actually increase its profit. In addition, we examine the importance of information provided by the platform, including the actually product quality and customer preference. Counter-intuitively, we show that retailer may not always use the perfect information as compliance with malicious customers can obtain positive ratings from them. This work also generates important implications for both retailers and platforms when dealing with malicious reviews.

Keywords: Online reviews, review manipulation, malicious consumers

1. Introduction

While most online reviews are based on genuine experiences, some reviews are intentionally distorted.[1] In this paper, we define malicious customers as those who make malicious complaints, which are made with the intention of causing harm through lying about an issue. Many cases have been reported that malicious customers contact retailers, ask for refunds, and threaten them with bad reviews if the retailers do not comply. Specifically, some malicious customers ask for financial compensation (e.g., refunds), pretending that they did not get the promised-quality product or service.[23] If the retailer does not comply with the requests, the malicious customers may post negative reviews, impacting the retailer's overall review rating.[4] If the retailer gives refunds, then there may be no negative reviews from malicious customers.

Because removing a distorted (or fake) negative review takes a long time or may not be allowed, many retailers face a tough decision to make. That is, the retailer should decide how to deal with such threats, i.e., whether to comply or not in a timely manner. It has been reported that many retailers just give refunds although they believe that malicious customers made false complaints.[5] Some restaurant owners even collapsed and died after handling complaints for full refunds from presumably malicious customers.[6] The existence of malicious customers creates quite serious problems for retailers. First, retailers may lose profits even though their product/service has no quality issues. Second, the retailers' overall review ratings might be underscored due to negative reviews from malicious customers.

In this paper, we investigate the issues caused by malicious customers and develop intuitions for the key questions related to the impact of malicious customers on retailers and regular consumers. Our first question is: What is the optimal response strategy for a retailer when she faces complaints from both malicious and regular (or good) customers and **cannot** tell who malicious customers are? We demonstrate that there exists a cutoff point for a retailer to choose the best strategy, either giving refunds to all customers who complained or turning down all refund requests. We show that both strategies would cause the retailer's overall ratings

[1] https://www.thecrimson.com/article/2022/3/29/online-food-reviews-and-yelp/
[2] https://www.latimes.com/food/story/2021-02-23/restaurant-scam-credit-card-fraud
[3] https://www.news10.com/news/local-restaurants-struggling-even-more-due-to-fraudulent-grubhub-orders/
[4] https://www.nj.com/news/2022/04/nj-restaurant-goes-viral-for-berating-threatening-customer-over-delivery-dispute.html
[5] https://www.winknews.com/2021/03/05/restaurants-hurt-by-customers-demanding-refunds-through-food-delivery-services/
[6] https://restofworld.org/2022/south-korea-star-ratings-trouble/

URI: https://hdl.handle.net/10125/106881
978-0-9981331-7-1
(CC BY-NC-ND 4.0)

to be distorted. In terms of consumer surplus, full compliance benefits regular consumers in general, while no-compliance strategy yields undesirable surplus loss from regular customers. Interestingly, we find that giving refunds to malicious customers could increase the retailer's profit due to distorted ratings.

Second, to investigate the platform's best strategy, we consider a case in which the platform could provide a mechanism that could track the actual quality of product or service provided. Using this model setup, we investigate the retailer's best response strategy. Counterintuitively, we find that the retailer may not choose to use the perfect mechanism because the retailer's profit could be higher by giving refunds or turning down refund requests to all customers who complained. This result implies a shocking managerial insight to practitioners. Even though most problems, if not all, caused by malicious customers can be appropriately handled if platforms could provide an effective quality tracking mechanism, our result shows that such a mechanism would not be used by retailers. This result highlights the complexity and importance of the research problem discussed in the paper.

2. Literature Review[7]

Online customer rating has been widely studied in the related fields of information systems and operations management. Online review systems provide an important channel for consumers to shared their purchase experience, and can further inform future consumers about the product quality and resolve concerns regarding product mismatch (Yu et al., 2016). Together with other information available on the platforms, such as product descriptions and sales volume, online reviews can also impact the strategies of both the firm and consumers (Liu et al., 2017; Yu et al., 2016).

In online retailing , the information a consumer can collect about the product or service is limited, thus the purchase decision can be heavily driven by the ratings/reviews available online (Li et al., 2011). In addition to product quality, the product fit information may not be available to consumers when they make their purchase decisions, however, online reviews provide "matched" or "unmatched" signals to consumers to resolve their perceived product value uncertainty to some extent (Li et al., 2011). Both quality and fit information can be learnt from reviews (Kwark et al., 2014), and sometime consumers are informed with a probability to know the true degree of misfit (or quality difference)

[7]Complete literature review is available upon request.

Consumers decisions are usually modeled in two periods (Jiang & Yang, 2019; Li & Hitt, 2010) where consumers arrive and purchase based on belief of expected quality and price, and consumers use reviews to form expectation of quality before purchase. Based on the reviews, consumers can update their belief on product valuation, which has been modeled as a weighted average between non-review-based belief and review-based belief in the literature (e.g., Y. Chen & Xie, 2008; Hao & Tan, 2019; Kwark et al., 2014; Yu et al., 2016). Consumers in the second period can obtain information about first-period demand, average rating, and variance of ratings (Sun, 2012). Moreover, the rating variance can be decomposed into taste-difference-caused variance and quality-difference-caused variance, and Zimmermann et al., 2018 have shown that the two types of rating variance matter in firm's pricing strategy under both consistent and inconsistent quality product marketplaces. As reviews could assist consumers learn the true valuation in the late period, the proportion of informed consumers in the second-period is usually related to the first-period demand (Jing, 2011).

As consumers' decisions heavily depend on the available review information, it will influence firm's decisions as well. When the quality is the firm's private information (signaling) and the online reviews could reveal the true quality level in the later period, the availability of online reviews will shift a firm's price and quality decisions (Jiang & Yang, 2019). In addition, different types of user-generated-content (e.g., rating, detailed review, information about utility or taste) are also found to influence advertising and pricing decisions (Zimmermann et al., 2018). Moreover, online reviews can intensify price competition and lead to lower profits for competing brands (Li et al., 2011). The impact of online customer reviews not only resides in retailer competition (or end-product market), which also extends to the upstream relationships in a channel network structure (Kwark et al., 2014).

Consumers' manipulation of online ratings has drawn significant attention in more recent work (Wu et al., 2020). Studies have shown that both extrinsic factors (e.g., monetary compensation, sense of mastery) and intrinsic factors (e.g., retaliation from an upset shopping experience) can motivate consumers to post negative fake reviews (Zaman et al., 2023). Contextual indicators, such as third-party labels and social proof, can assist potential customers in their efforts to detect fake reviews (Munzel, 2015). However, most of the current work focuses on the detection of fake reviews (e.g., Fayazi et al., 2015) or explanation of consumers' incentives (e.g., Zaman et al., 2023), there is a lack of understanding how the phenomenon influences the platform and retailers' profits and what the optimal strategic response could be.

3. Model Formulation and Benchmarks

3.1. Consumer Purchase Decisions

We consider a model where a retailer provides its product (or service) through an online platform. A customer who purchased the product may not always receive the product at the expected quality due to the inherent process variability of the retailer and/or the imperfect delivery process. To represent this in a parsimonious manner, we assume the product quality received by the customer Q as follows:

$$Q = \begin{cases} 1 & \text{non-defective with the probability of } \tilde{\alpha}, \\ 0 & \text{defective with the probability of } 1 - \tilde{\alpha}, \end{cases}$$

where $\tilde{\alpha}$ is the true probability that a customer receives a good product (service). While the true probability $\tilde{\alpha}$ is not known to customers, customers can observe the ratings on the platform, which were evaluated by previous customers. We model that consumers differ in their preferences denoted by x and uniformly distributed between 0 and 1, i.e., $x \in U[0, 1]$. We also introduce the intrinsic value of the product (service), v, which is not subject to the product quality. Then, the consumer's utility can be expressed as follows:

$$u(x|Q) = v + Q - mx - p, \tag{1}$$

where m denotes the misfit or travel cost and p is the price for the product (service).

Since the product quality is unknown at the purchase, the consumer's expected utility is $U(x) = v + \alpha_t Q - mx - p$. Then, the marginal consumer in period t is located at $\bar{x}_t = \frac{\alpha_t + v - p}{m}$. As we assume the potential market size is one unit in each time period, the demand for the product (service) in period t is $D_t = \bar{x}_t$.

3.2. Two Customer Types and Their Rating Processes

Since the product quality is binary, it is reasonable to model that the customer's experience is either satisfied or dissatisfied (Papanastasiou et al., 2023). For ease of exposition, we consider the consumer's satisfaction rate as the online rating, which is the proportion of positive ratings among all consumers (J. Chen et al., 2022). However, as noted earlier, not all customers rate the retailer based on the true quality they observe.

We let $\beta \in [0, 1]$ be the proportion of malicious customers, whose rating is contingent on the retailer's response rather than the product quality. In contrast, there is the $1 - \beta$ proportion of regular customers, whose ratings always depend on the actual product quality and after-sales service they have received. To present clear insights, we model that all customers will post ratings for the retailers after purchase.

Suppose that a regular customer receives the product and observes the product quality. If the product is good ($Q = 1$), her utility is $u(x|1) = v + 1 - mx - p > 0$ and gives a positive rating. If the product is defective ($Q = 0$), we assume that she contacts the retailer and requests a full refund. When the retailer does not comply with the refund request, she proceeds and posts a negative rating since $u(x|0) = v - mx - p < 0$. However, if the retailer accepts her request, the rating depends on her utilities, $u(x|0) + p = v - mx$. If $u(x|0) + p > 0$ or equivalently $x < \bar{x}' = v/m$, she rates the retailer positively. Otherwise, she gives a negative rating. Therefore, $\bar{x}' = v/m$ proportion of consumers who received defective products still rate the retailer positively after getting a full refund, but the rest $\bar{x}_t - \bar{x}'$ will rate negatively. Figure 1(a) illustrates the regular customer's rating process.

The other type of customers is malicious in that they always argue that their products are defective regardless of the actual quality and ask for refunds. We assume that the malicious customers would post negative reviews if the retailer did not comply with their requests. To make their threats more appealing, they also propose positive ratings if the retailer does comply. Figure 1(b) shows the malicious customer's rating process.

3.3. Online Ratings and the Retailer's Demand

We suppose that the retailer should determine its internal compliance policy, which will stand for a period of time. In practice, this period can be months, quarters, or years. To capture the impact of current decision on the foreseeable future in a concise manner, we study a two-period model. As discussed above, the demand is contingent on not only price and consumers' preference but also the customer's rating, which is affected by the firm's compliance policy.

At the beginning of the first period ($t = 1$), the cumulative product rating is α_1. Customers arriving in the first period will make their purchase decisions based on the observed prior ratings and the product price, which leads to the marginal customer located at $\bar{x}_1 = \frac{\alpha_1 + v - p}{m}$ and the demand for the retailer is $D_1 = \bar{x}_1$.

Among all the D_1 customers in the first period,

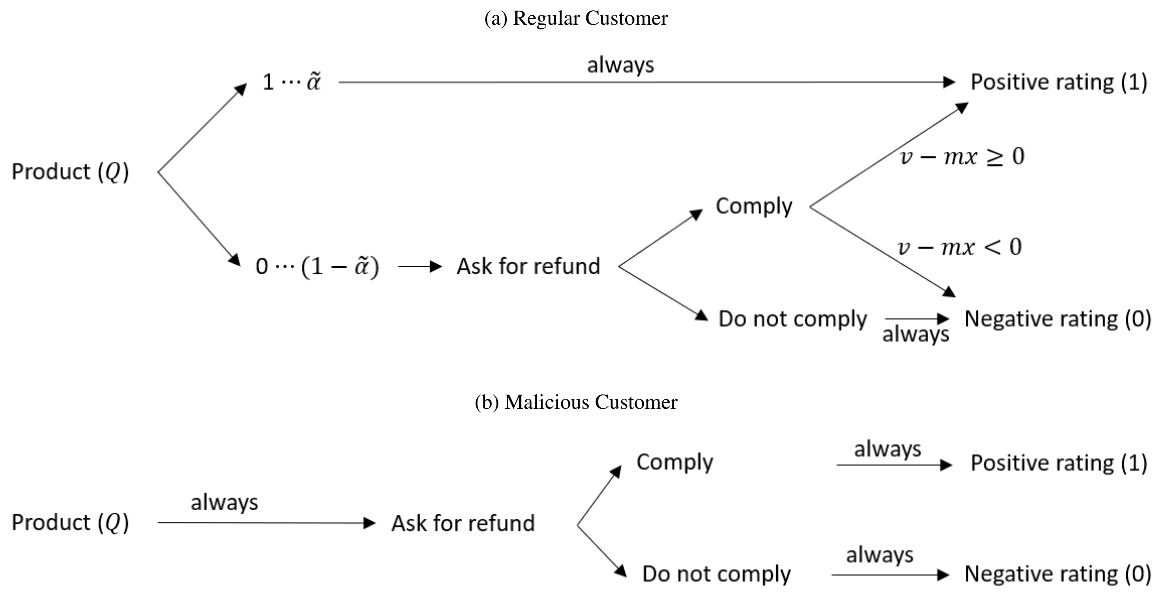

Figure 1. Customer's Rating Process

there are β proportion malicious customers who always request refunds and the rest $(1-\beta)$ are regular customers who request refunds if they received the defectives. Anticipating this, the retailer should determine the compliance policy. In the baseline model, it is reasonable to assume that the retailer cannot distinguish the customer's type (regular vs. malicious) or consumer preferences. Thus, the retailer has a consistent policy all refund requests. That is, when the retailer cannot distinguish the customer's type, the retailer should either accept or reject all refund requests. As a result, the retailer's compliance policy is either *Full Compliance* (F) or *No Compliance* (N).

When the retailer adopts F, the total number of accepted refund requests (R) is the sum of the number of malicious customers ($D_1\beta$) and the regular consumers who receive the defectives ($D_1(1-\beta)(1-\tilde{\alpha})$). For the no compliance case, R is just zero. That is, R at the beginning of the first period for each policy can be expressed as follows:

$$R_{i,1} = \begin{cases} D_1(\beta + (1-\beta)(1-\tilde{\alpha})) & \text{for } i = F, \\ 0 & \text{for } i = N. \end{cases}$$

At the end of the first period, customers make their ratings decisions based on the product/service they have received and the compliance policy adopted by the retailer according to the processes in Section 3.2. Then, the accumulated ratings received by the seller get updated, which becomes the prior information that influences the purchase decisions of customers arriving in the second period. The literature has documented the influence of average product ratings on consumers' decision-makings (Li & Hitt, 2008; Li, 2018). Therefore, the cumulative online rating at the beginning of the second period ($t = 2$) can be updated as

$$\alpha_{i,2} = (1-\lambda)\alpha_1 + \lambda$$
$$\begin{cases} [(1-\beta)(\tilde{\alpha}\bar{x}_1 + (1-\tilde{\alpha})\bar{x}') + \beta\bar{x}']/D_1 & \text{for } i = F, \\ \tilde{\alpha}(1-\beta) & \text{for } i = N, \end{cases}$$

in which λ represents the relative weight of recent reviews on the overall rating score.

Based on the updated accumulative ratings $\alpha_{i,2}$, the retailer's demand for the second period can be generally expressed as

$$D_{i,2} = \bar{x}_{i,2} = \frac{\alpha_{i,2} + v - p}{m}, \qquad (2)$$

where $i \in \{F, N\}$ denotes the compliance policy. We note that as the compliance policy is internal and determined at $t = 1$, D_1 is independent of i but $D_{i,2}$ is contingent on i. As the compliance policy adopted at the beginning of the first period remains unchanged in the second period, the number of refunds issued in the second period can be derived as

$$R_{i,2} = \begin{cases} D_{i,2}(\beta + (1-\beta)(1-\tilde{\alpha})) & \text{for } i = F, \\ 0 & \text{for } i = N. \end{cases}$$

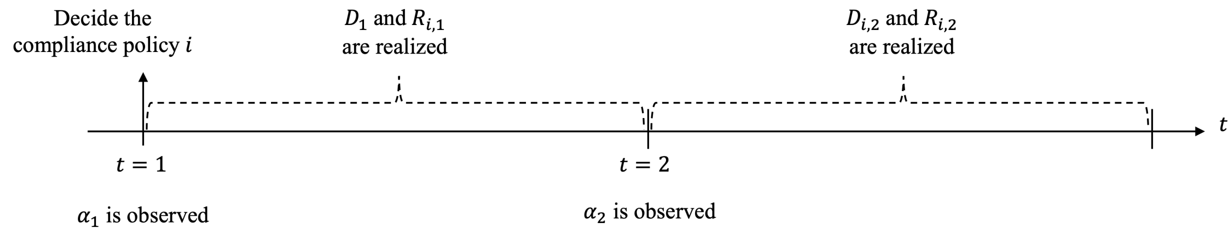

Figure 2. The Sequence of Events

Therefore, the retailer objective functions can be expressed as

$$\max_i \pi(p, i) = p(D_1 - R_{i,1}) + p(D_{i,2} - R_{i,2}). \quad (3)$$

We summarize the sequence of the events in Figure 2. First, at the beginning of the first period, the retailer decides the compliance policy (i) for the cumulative product ratings (α_1). Second, during the first period, the demand and the refund requests are realized. The retailer deal with the requests according to its compliance policy. Third, at the beginning of the second period, the ratings are updated by customers from the previous period. The demand and the refund requests realizations are repeated.

3.4. Benchmark Case: When No Information is Available

Proposition 1. *(i) The retailer adopts Full Compliance (F) if its initial rating α_1 is small enough, i.e., $\alpha_1 \leq \alpha_F^P(\alpha_1, \beta)$. Otherwise, the retailer adopts No Compliance (N).*

(ii) Moreover, the retailer is more likely to adopt F when its initial rating is low and there are more malicious consumers, i.e., $\alpha_F^P(\alpha_1, \beta)$ is decreasing in α_1 but increasing in β.

Proposition 1 shows that the retailer will comply more when it is under harsh business conditions, namely, low quality ($\tilde{\alpha}$), low rating (α_1), and more malicious consumers (β). In contrast, if the retailer is under favorable conditions, it does not comply at all.

An interesting question is whether the retailer can earn more profit with malicious consumers. In particular, what if the retailer can sell only to a smaller population of regular consumers ($1 - \beta$) instead of the whole population including the malicious consumers? Will the retailer prefer this exclusive market? Or, will it still want to have a larger and more inclusive market even with malicious consumers?

Proposition 2. *For some proportion of malicious consumers β, if the initial rating is high enough, the retailer's profit is greater with β of malicious consumers and $1-\beta$ of regular consumers than only $1-\beta$ of regular consumers.*

Proposition 2 shows that the retailer largely prefers not to have malicious consumers. Recall that the firm can earn a higher rating at beginning of the second period in exchange for profits via compliance. Malicious consumers have the retailer forego more profits for the same amount of positive reviews. As such, if the initial rating is low to moderately high, the retailer will be better off without malicious consumers. Nevertheless, when the initial rating is high enough, it is indeed better off in a larger market with malicious consumers. As Proposition 1 shows, the retailer will adopt N, not complying with any complaints. Even if such a strict policy lowers its second-period rating, the negative impact is mitigated thanks to the high initial rating. Therefore, the retailer benefits from selling to malicious consumers but does not comply with their refund requests, as illustrated in the shaded region in Figure 3(b).

Next, we examine under which policy consumers are better off. When the retailer adopts F, the consumer having the defective product receives a full refund. However, we remind that she still incurs negative utilities because of preference misfit (tx). Since the second-period rating and demand are higher than those under N, there will be more refunds, potentially resulting in a larger aggregated misfit cost. We compare the consumer's surplus under each policy in the next proposition.

Proposition 3. *Consumers are always better off under full compliance (F) than no compliance (N).*

Despite the aforementioned trade-off, the consumers as a whole obtain more surplus as the retailer is more compliant. In conjunction with Proposition 1, this result implies that the regular consumers can be indeed better off when there are more malicious consumers because the retailer's compliance policy may change.

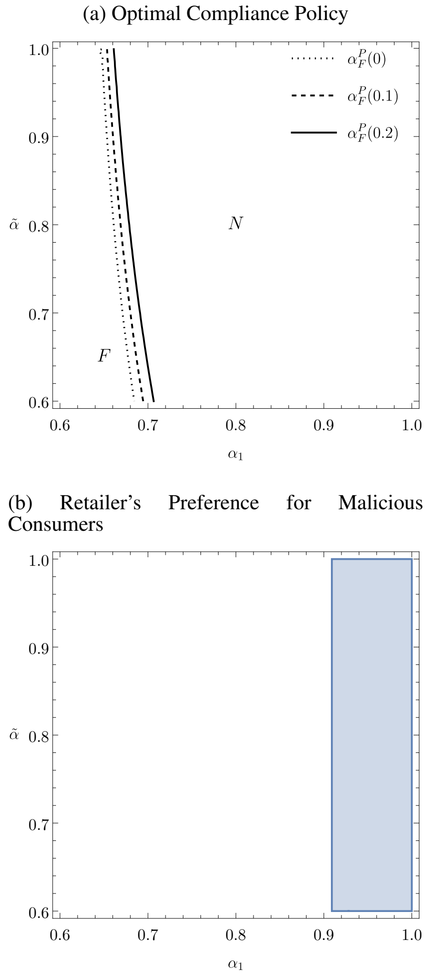

Figure 3. The Impacts of Malicious Consumers on the Retailer without any Information

4. Information Provided by the Platform

4.1. When Actual Quality Information is Available

As the platform is an important agent in the transaction, it usually records detailed information regarding each transaction. For instance, a delivery company may request drivers to take a picture of the packages delivered. Given the presence of malicious consumers, the platform is able to scrutinize every transaction and identify whether the refund is requested by a customer who does receive a defective product or a malicious customer who actually receives non-defective product. Although such identification may never be perfectly accurate, the accuracy could be reasonably high due to the availability of data collected through IoT devices and the advancement of detection algorithms. We assume that the actual quality of product (or service) received by customers is available and provided to the retailer.

There are a few different ways that the retailer can utilize the actual product quality information. On the one hand, it can still fully (F) or never (N) comply, practically not utilizing the information at all. On the other hand, it only complies with the complaints from consumers that did receive defective products, which is referred to as Conditional Compliance (C). That is, if any consumer receives a defective, then the retailer complies with the refund request. However, if a malicious consumer with a good product pretends to have a defective and requests a refund, the retailer does not comply with the request despite the negative rating from him.

Under C, the consumer utility and market demand in the first period are the same as the baseline cases. But considering all requests from consumers, the retailer will only refund customers who actually received defective products:

$$R_{C,1} = D_1(1-\tilde{\alpha}).$$

Then, consumers rate their purchase experience based on the retailer's response and the accumulated ratings received by the seller get updated. As the retailer turns down all requests from malicious consumers who receive non-defective products and only refunds customers who receive defective ones, the cumulative online rating at the beginning of the second period ($t = 2$) can be updated as

$$\begin{aligned}\alpha_{C,2} =& (1-\lambda)\alpha_1 + \lambda[(1-\beta)(\tilde{\alpha}\bar{x}_1 + (1-\tilde{\alpha})\bar{x}') \\ &+ \beta(1-\tilde{\alpha})\bar{x}']/D_1.\end{aligned}$$

As the compliance policy adopted at the beginning of the first period remains unchanged in the second period, the number of refunds issued in the second period can be derived as

$$R_{C,2} = D_2(1-\tilde{\alpha}).$$

Since all regular consumers' refund requests will be accepted, the consumers' surplus for the conditional comply case is the same as that for the fully comply case, which is denoted as:

$$\begin{aligned}CS_{C,t} = \tilde{\alpha}\int_0^{\bar{x}_{C,t}} (v+1-mx-p)\,dx \\ + (1-\tilde{\alpha})\int_0^{\bar{x}_{C,t}} (v-mx)\,dx.\end{aligned}$$

Then, how does the retailer's optimal policy change?

Proposition 4. *Suppose that the actual product quality information is provided to the retailer.*

(i) The retailer adopts Conditional Compliance (C) if its initial rating α_1 is small enough, i.e., $\alpha_1 \leq \alpha^P_{CN}(\alpha_1, \beta)$. Otherwise, the retailer adopts No Compliance (N).

(ii) Moreover, the retailer is more likely to adopt C when its initial rating is low and there are more malicious consumers, i.e., $\alpha^P_F(\alpha_1, \beta)$ is decreasing in α_1 but increasing in β.

(iii) The retailer adopts C if its quality is high enough regardless of its initial rating.

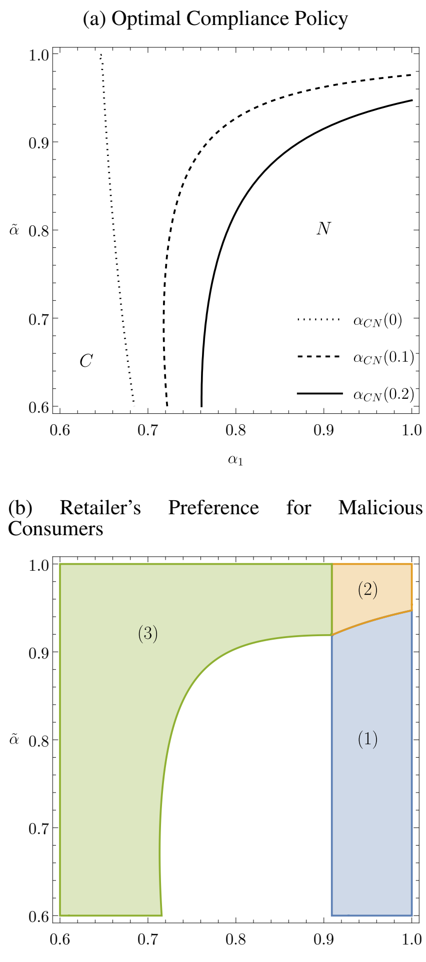

Figure 4. The Impacts of Malicious Consumers on the Retailer with Actual Product Quality Information

As the actual product (or service) quality information enables the retailer to respond to the consumers' complaints more flexibly, the retailer is more likely to prefer a larger market with malicious consumers. Figure 4(b) illustrates this. The combination of Regions (1) and (2) is identical to the shaded region in Figure 3(b). However, Region (2) indicates that the retailer would be even more beneficial with C. Region (3) exhibits the additional region that the retailer prefers a larger market thanks to the type information.

4.2. When Consumer Preference is Available

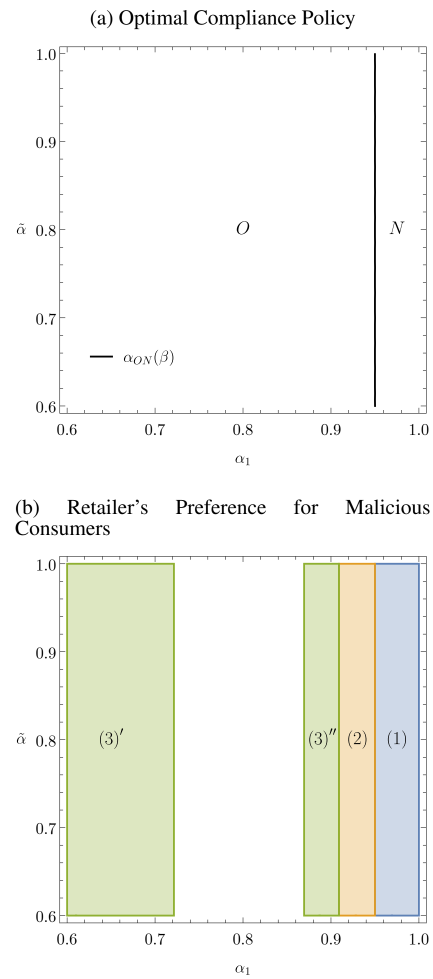

Figure 5. The Impacts of Malicious Consumers on the Retailer with Consumer's Preference Information

Because of consumer preference or the heterogeneous misfit cost, those who have a larger misfit ($x > \bar{x}'$) still suffer from utility losses even receiving a refund from the retailer. In other words, the full refund practice may not guarantee perfect customer satisfaction. For the retailer's rating concern, such preference information is critical. Suppose that the platform provides the consumer's preference information so that the retailer can decide whether to offer a refund based on customers' potential ratings. This is regarded as the retailer can opportunistically comply with those who would give a positive rating, which is referred to as O. Put differently, the retailer efficiently *buys* positive ratings for a deceitful second-period rating.

In the first period, among regular consumers $(1-\beta)$ who purchased $(\bar{x}_1)$ and received defective products $(1-\tilde{\alpha})$, the $\bar{x}_1 - \bar{x}'$ proportion still gets negative utility after receiving a full refund. Among malicious consumers (β) that all claim having received defective products, because the retailer cannot tell if the customer actually received defective or non-defective product, the retailer is assumed to comply when the customer is located between $[0, \bar{x}']$. As a result, the refund provided by the retailer in the first period is:

$$R_{O,1} = (1-\beta)(1-\tilde{\alpha})\bar{x}' + \beta\bar{x}'.$$

As these customers will post positive ratings, the cumulative ratings are updated as follows:

$$\begin{aligned}\alpha_{O,2} =& (1-\lambda)\alpha_1 \\ &+ \lambda[(1-\beta)(\tilde{\alpha}\bar{x}_t + (1-\tilde{\alpha})\bar{x}') + \beta\bar{x}']/D_1.\end{aligned}$$

Again, as the compliance policy in the second period remains the same, the amount of refund provided by the retailer is

$$R_{O,2} = (1-\beta)(1-\tilde{\alpha})\bar{x}' + \beta\bar{x}'.$$

In the opportunistic case, not all regular customers' refund requests are accepted. Regular customers that received defective products can only get a full refund if they are located in $[0, \bar{x}'_{O,t}]$. For the rest (located in the interval $[\bar{x}'_{O,t}, \bar{x}_{O,t}]$), the retailer will not consider their refund request. Thus, the consumer surplus differs from the fully comply or conditional comply case.

$$\begin{aligned}CS_{O,t} =& \tilde{\alpha}\int_0^{\bar{x}_{O,t}} (v+1-mx-p)\,dx \\ &+ (1-\tilde{\alpha})[\int_0^{\bar{x}'_{O,t}} (v-mx)\,dx \\ &+ \int_{\bar{x}'_{O,t}}^{\bar{x}_{O,t}} (v-mx-p)\,dx].\end{aligned}$$

Figure 5(a) illustrates that the opportunistic policy is primarily optimal unless the initial rating is high enough. Similar to when the consumer type information is available, the retailer will not fully comply with the complaints when the preference information is available. When the retailer is compliant, it utilizes the available information and adopts O as its initial rating is lower. We note that the indifferent line is independent of β. It is because the retailer's profit difference between N and O is how much refund it offers. Recall that the size of those who leave a positive rating given refund is $\bar{x}'$, invariant to β. So is the profit difference.

Figure 5(b) exhibits somewhat similar features to Figure 4(b). The combination of Regions (1) and (2) is identical to the shaded region in Figure 3(b). Region (2) is where the retailer can be more beneficial by adopting C instead of N. Both Regions $(3)'$ and $(3)''$ are an additional area compared to the no information case.

4.3. When Both Information is Available

Consider that the platform can decide which type of information to provide to the retailer. The adoptable compliance policies are F, N, C, and O. As discussed before, F is not optimal when any information is available. Among the other three alternatives, Figure 6 illustrates that the retailer adopts O unless either the product quality or the initial rating is high enough. When the quality is high enough, the retailer can adopt C and comply with the true complaints. However, if the quality is not excellent but the initial rating is high, ignoring all complaints (N) is optimal. Furthermore, Figures 6(a) and 6(b) show that C is more likely optimal as there are more malicious consumers. This implies that if the platform with more malicious consumers should choose to invest in improving the accuracy of one kind of information, it needs to focus more on the consumer type information.

5. Conclusion

Malicious consumers on a platform, driven by ill intent rather than genuine feedback, present a unique challenge for the platform and its retailers. These consumers often exploit the unverifiability of actual product quality by making unfair claims and demanding refunds or replacements. If the retailer refuses to comply, they retaliate by leaving negative reviews. Given that online reviews significantly influence consumer purchasing decisions, malicious consumers can sway potential customers away from the retailer's products or services. Balancing the need to address legitimate claims while tackling false or malicious claims becomes crucial for retailers in maintaining their reputation on the platform. To assist retailers in managing this challenge, platforms may allocate significant resources to monitor user activities, develop algorithms, and provide information to retailers. This paper studies the implications of the availability

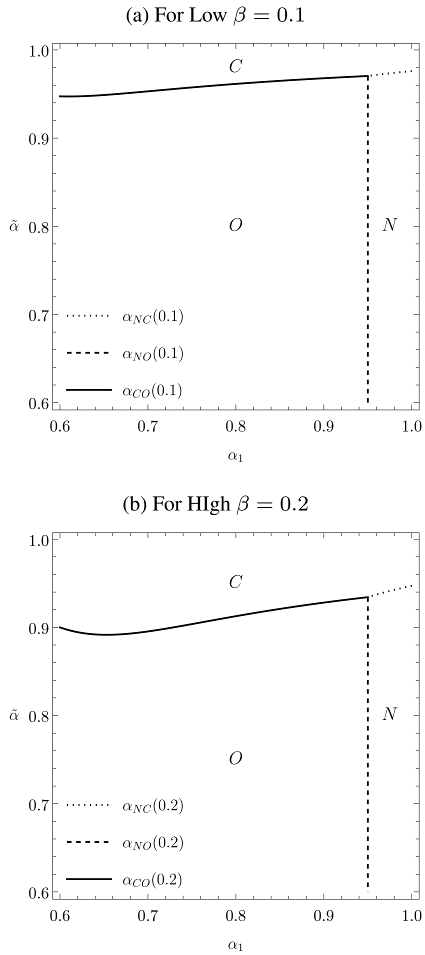

Figure 6. The Retailer's Optimal Policy with Actual Product Quality and Consumer Preference Information

of different types of information on the retailer's compliance policy and consumer welfare.

This study contributes to the literature by investigating the influence of malicious consumers on retailers' compliance responses. Although the phenomenon of online rating and retailer's review manipulation strategies have been widely studied, the existence of malicious consumers and its potential impact are not well-understood. This study bridges this gap by examining the retailer's optimal response facing the threat by malicious users and demonstrating the influence on regular consumers' surplus. This work also generates important implications for both retailers and platforms. Mitigating the pain caused by malicious users on retailers requires a collaborative effort. The platform should strive to provide valuable information to retailers, and retailers, in turn, should implement it fairly, striking a balance between profit and consumer surplus. By achieving this balance, the platform and retailers can create a fair and secure marketplace while minimizing any potential negative impact on regular consumers in the presence of malicious consumers.

There are a few future directions to extend our research. First, we assumed that malicious consumers would always provide positive ratings if a refund request is granted by retailers. While we believe that relaxing this assumption to proportional positive ratings (i.e., some malicious consumers will give positive ratings and others will not provide any ratings) would not change our main findings significantly, further investigation is needed to see the changes of various cutoff points for the optimal strategy. Second, investigating the impact of product prices on the optimal strategy might yield some interesting managerial insights because firms selling high-price products should have different strategies compared to those selling low-price products.

References

Chen, J., Guo, Z., & Huang, J. (2022). An economic analysis of rebates conditional on positive reviews. *Information Systems Research*, *33*(1), 224–243.

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, *54*(3), 477–491.

Fayazi, A., Lee, K., Caverlee, J., & Squicciarini, A. (2015). Uncovering crowdsourced manipulation of online reviews. *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*, 233–242.

Hao, L., & Tan, Y. (2019). Who wants consumers to be informed? facilitating information disclosure in a distribution channel. *Information Systems Research*, *30*(1), 34–49.

Jiang, B., & Yang, B. (2019). Quality and pricing decisions in a market with consumer information sharing. *Management Science*, *65*(1), 272–285.

Jing, B. (2011). Social learning and dynamic pricing of durable goods. *Marketing Science*, *30*(5), 851–865.

Kwark, Y., Chen, J., & Raghunathan, S. (2014). Online product reviews: Implications for retailers and competing manufacturers. *Information systems research*, *25*(1), 93–110.

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456–474.

Li, X., & Hitt, L. M. (2010). Price effects in online product reviews: An analytical model and empirical analysis. *MIS quarterly*, 809–831.

Li, X., Hitt, L. M., & Zhang, Z. J. (2011). Product reviews and competition in markets for repeat purchase products. *Journal of Management Information Systems*, *27*(4), 9–42.

Li, X. (2018). Impact of average rating on social media endorsement: The moderating role of rating dispersion and discount threshold. *Information Systems Research*, *29*(3), 739–754.

Liu, Y., Feng, J., & Liao, X. (2017). When online reviews meet sales volume information: Is more or accurate information always better? *Information Systems Research*, *28*(4), 723–743.

Munzel, A. (2015). Malicious practice of fake reviews: Experimental insight into the potential of contextual indicators in assisting consumers to detect deceptive opinion spam. *Recherche et Applications en Marketing (English Edition)*, *30*(4), 24–50.

Papanastasiou, Y., Yang, S. A., & Zhang, A. H. (2023). Improving dispute resolution in two-sided platforms: The case of review blackmail. *Management Science*.

Sun, M. (2012). How does the variance of product ratings matter? *Management Science*, *58*(4), 696–707.

Wu, Y., Ngai, E. W., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, *132*, 113280.

Yu, M., Debo, L., & Kapuscinski, R. (2016). Strategic waiting for consumer-generated quality information: Dynamic pricing of new experience goods. *Management Science*, *62*(2), 410–435.

Zaman, M., Vo-Thanh, T., Nguyen, C. T., Hasan, R., Akter, S., Mariani, M., & Hikkerova, L. (2023). Motives for posting fake reviews: Evidence from a cross-cultural comparison. *Journal of Business Research*, *154*, 113359.

Zimmermann, S., Herrmann, P., Kundisch, D., & Nault, B. R. (2018). Decomposing the variance of consumer ratings and the impact on price and demand. *Information Systems Research*, *29*(4), 984–1002.