



Information Systems Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Measuring and Managing the Externality of Managerial Responses to Online Customer Reviews

Wei Chen, Bin Gu, Qiang Ye, Kevin Xiaoguo Zhu

To cite this article:

Wei Chen, Bin Gu, Qiang Ye, Kevin Xiaoguo Zhu (2019) Measuring and Managing the Externality of Managerial Responses to Online Customer Reviews. Information Systems Research

Published online in Articles in Advance 29 Jan 2019

. <https://doi.org/10.1287/isre.2018.0781>

Full terms and conditions of use: <https://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2019, INFORMS

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.



For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Measuring and Managing the Externality of Managerial Responses to Online Customer Reviews

Wei Chen,^a Bin Gu,^b Qiang Ye,^{c,*} Kevin Xiaoguo Zhu^d

^a Eller College of Management, University of Arizona, Tucson, Arizona 85721; ^b W. P. Carey School of Business, Arizona State University, Tempe, Arizona 85287; ^c School of Management, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China; ^d Rady School of Management, University of California, San Diego, California 92093

* Corresponding author

Contact: weichen@email.arizona.edu,  <http://orcid.org/0000-0002-0963-7839> (WC); bin.gu@asu.edu (BG); yeqiang@hit.edu.cn,  <http://orcid.org/0000-0002-0963-7839> (QY); kxzh@ucsd.edu (KXZ)

Received: December 17, 2015

Revised: February 2, 2017; December 1, 2017

Accepted: January 14, 2018

Published Online in Articles in Advance:
January 29, 2019

<https://doi.org/10.1287/isre.2018.0781>

Copyright: © 2019 INFORMS

Abstract. Managerial responses to online customer reviews not only affect customers who receive the responses but may also influence subsequent customers who observe the responses. This externality arises because of the public nature of online interactions. However, prior studies were mainly in offline settings where such externality rarely exists. In this study, we assess the magnitude of such externality. Using a difference-in-differences framework and matched hotels across two large travel agencies, we find that managerial responses indeed have a significant and positive impact on the *volume* of subsequent customer reviews. The impact on the review *valence* is not evident, which can be attributed to the unique design of identity disclosure in our research context. Furthermore, our results suggest nuances that were not known in the prior literature. For example, responding to positive and negative reviews may have different effects on future reviews, and managers should provide detailed responses to negative reviews but brief ones to positive reviews. Our results offer managerial implications to service providers on how to improve customer engagement in the interconnected online environment.

History: Kai-Lung Hui, Senior Editor; ParamVir Singh, Associate Editor.

Funding: This research is partially supported by the National Science Foundation of China [Grants 71225003, 71490724, 71532004, 71620107005, and 71328102] and 111 project [Grant B18043].

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/isre.2018.0781>.

Keywords: online interactions with customers • managerial responses • customer reviews • externality • volume • valence • mechanisms

1. Introduction

Over the last decade, online customer reviews have become an important source of product information for consumers. More than 80% of online shoppers consult reviews by other customers before making purchase decisions (Anderson 2014), and consumers view such reviews as more neutral or trustworthy than information provided by retailers (Bickart and Schindler 2001). Previous research shows that customer reviews have a significant influence on product sales (see, e.g., Chevalier and Mayzlin 2006, Chintagunta et al. 2010, Duan et al. 2008b).

Given the importance of customer reviews, businesses increasingly take a proactive approach to managing customer reviews. First, companies actively solicit product assessments from reviewers. For example, the Amazon Vine program provides vendor-sponsored complimentary products to a select group of reviewers so they can “post opinions about new and prerelease items to help their fellow customers make educated purchase decisions.”¹ Another common (yet controversial) practice is censoring, though

the effectiveness of censoring is questionable. Moreover, censoring is applicable only to reviews hosted by the retailer. Consumers now increasingly turn to third-party websites, which are believed to be less biased. Third, companies have recently begun to post promotional reviews on third-party review sites (Mayzlin et al. 2014) despite the fact that this approach is costly for both the business and the review platforms. As a result, review platforms spend great efforts to detect fake reviews. Businesses may also suffer a backlash if customers discover phony positive reviews.

In this study, we explore an alternative approach that has been underexplored in the literature: online managerial responses (MRs). We note that many customer review platforms (e.g., TripAdvisor.com) allow businesses to submit MRs to consumer reviews. Some companies have started paying attention to MRs (Park and Allen 2013). An extensive survey of online review sites, however, indicates that the use of MRs is still limited despite the growing amount of customer reviews online (Lappas et al. 2016, Levy et al. 2013), for businesses seem uncertain about the pros and cons of MRs. The goal of this study is to assess the effectiveness

of MRs on future customer reviews and to investigate possible underlying mechanisms for the influence.

A unique feature of this study is that we leverage the *same* hotels on two major online travel agencies to address the endogeneity issue caused by the unobserved underlying management expertise of hotels. For example, the introduction of a new management team can simultaneously improve responses to online consumer reviews and the offline service quality. Both travel agencies in our context host extensive user reviews of hotels with one key difference: one travel agency allowed hotels to post responses to customer reviews, whereas the other did not (during our study period). Since we observe the same hotel across two sites, it is possible to control the influence of unobserved managerial expertise and isolate the influence of MRs by comparing consumer reviews of the same hotel across the two travel agencies before and after the provision of MRs.

Using a difference-in-difference-in-differences (DDD) approach and panel data from two large online travel agencies, we find that MRs have a significant impact on subsequent customer reviews. First, a hotel would receive about 12% to 14% more volume of reviews after it provides MRs. This increase in volume highlights the importance of MRs to online customer reviews. Second, we find a statistically nonsignificant impact on the valence of reviews, which may be because of the unique design feature of limited identity disclosure in our research context. Third, our further analysis of the time effect reveals that the influence of MRs is not permanent. Instead, if the vendor posts no more MRs after the first or previous ones, the effect on review volume will decay over time. Analysis using the number of MRs also confirms the results. Fourth, MRs to positive and negative reviews may have different effects on future reviews. Furthermore, detailed MRs that focus on the exact issues of reviewer's comments may be essential for negative reviews, but may attenuate the effect of responses to positive reviews.

These findings indicate that, in online business circumstance, MR is a very important new tool besides traditional customer recovery practices in handling online reviews and conducting customer recovery. Many online platforms, such as Amazon.com, Yelp.com, Newegg.com, and the Apple app store, have established functions for businesses to provide MRs to customer reviews. Even though MR is still not well employed by all firms, it shows the potential to be integrated into routine business procedures in the future, especially for business-to-consumer firms in service sectors. Theoretically, MR extends existing customer recovery theories into the more interconnected online environment. Different from previous theories, the open characteristics of online environments bring about the externality that we reveal in this study.

2. Literature Review

The impact of online reviews on sales has been widely documented in the literature (e.g., Chevalier and Mayzlin 2006, Duan et al. 2008a, Zhu and Zhang 2010), which calls for the careful attention of management. However, it is still unclear how business can influence the opinions of customers or the public at large, without alienating them. Recently, different disciplines have turned to the antecedents of online reviews (Berger and Schwartz 2011, Burtch et al. 2018). In this section, we first review the literature on drivers of online reviews, and then discuss how MR can drive online reviews through different mechanisms.

2.1. Drivers of Online Reviews

The *volume* of online reviews can be driven both by the underlying sales and reviewers' psychological motivations. Intuitively, the underlying sales may be positively associated with the volume of online reviews. Besides the large literature around the impact of reviews on sales (Chevalier and Mayzlin 2006, Ye et al. 2009), prior literature has also shown the correlation between sales and future review activities (Dellarocas and Narayan 2006, Ren et al. 2018). Because online reviews reflect customers' postpurchase evaluation of a product or service (Moe and Schweidel 2012), assuming that customers have a constant probability of posting reviews, the number of reviews should be positively associated with the underlying sales (Ye et al. 2011).

While the underlying sales form the basis for review *volume*, customers' propensity to post reviews can be affected by their own experience and various environmental factors (Moe and Schweidel 2012). Regarding individual experience, research has shown that customers with extreme opinions will be more likely to contribute online reviews (Dellarocas and Narayan 2006). Prior research has also suggested that disconfirmation of expectations, which are formed on top of previous reviews, may increase customers' motivation to write a review (Ho et al. 2017). Moreover, the social influence from the community may affect the review motivation. Previous studies show that a larger audience or more followers drive a reviewer to post more comments (Goes et al. 2014, Wang et al. 2017). The integration with social media platforms (e.g., Facebook) may also increase the review volume because of higher social presence (Huang et al. 2017). The social influence can further work together with financial incentives to stimulate review generation (Burtch et al. 2018).

The drivers of review *valence* can be classified into two categories: experience based and identity based. *Experience-based drivers* emphasize the unique experience of the customer who posts the review. For example, previous studies find that the average rating of online reviews tends to decline over time. This decline may come from that early buyers like the product

more (Li and Hitt 2008), or because people are less likely to post additional positive reviews for already highly rated products (Godes and Silva 2012, Wu and Huberman 2008). Research has also suggested that users contribute more negative reviews when they become more experienced (Moe and Schweidel 2012).

Identity-based drivers focus on the social influence of the community on the reviewer's willingness to post positive or negative reviews. Online review sites can be considered as a special case of online communities, especially when more and more such sites add social components (Goes et al. 2014). In the online community literature, community interactions may influence the contribution of users through their identities in the community (Chen et al. 2018). In the context of online reviews, identity disclosure shapes community members' judgment of reviews (Forman et al. 2008). Research has shown that social media integration may increase the review valence because people are afraid of social disapproval (Huang et al. 2017), and a larger number of followers may lead to lower ratings by reviewers (Goes et al. 2014). Reviewers may also exhibit differentiation or herding behaviors depending on whether they are influenced by strangers or friends (Lee et al. 2015).

The aforementioned literature reveals the mechanisms through which underlying sales, customer experience, and social influence can affect review volume and valence. Our study examines how reviewers' incentives and posting behaviors may change when they see the MRs of a vendor (i.e., a hotel). The results add to the literature by examining the possible mechanisms of MR, and providing suggestions for managerial response practices.

2.2. Mechanisms of Managerial Response

Even though hotels have begun to pay attention to MRs, only a small portion of hotels are using it (Levy et al. 2013). A recent study shows that around two-thirds of all negative reviews on TripAdvisor do not receive responses from businesses (Lappas et al. 2016). Also, hotels have diverse practices in using MRs (Park and Allen 2013). In this paper, we attempt to advance our understanding of whether the public nature of MRs in the online review environment would influence the review behaviors of future customers, and the possible underlying mechanisms for effective MRs.

MRs may affect future postings by altering the motivation to post reviews. The impact of responses on user motivation and contributions is evident in the online community literature. Community responses may sustain users' continued participation in open-source software (Zhang et al. 2013) or knowledge sharing communities (Chen et al. 2018). Firms' information content on social media may also affect consumers' engagement and content generation on Facebook (Lee et al. 2018). Regarding MRs, Proserpio and Zervas (2017) have a

similar setting to ours, and find that the adoption of MR increases both review volume and valence. They argue that MRs encourage consumers to post more positive reviews but fewer negative reviews, because consumers who post positive reviews may feel that their feedback is appreciated, while those who post negative reviews know that their feedback will be scrutinized. Because this explanation focuses on the motivation to post reviews, we call it the *motivation* argument.

We add to this literature by arguing that a design feature of online review sites—the disclosure of identity information—may affect the influence of MRs on review valence. Specifically, we find a similar effect of MRs on review volume as in the literature, but we do not find a significant increase in valence after MR adoption. As we have reviewed in Section 2.1, identity-based drivers can affect the willingness to post reviews through the social influence of the community (Forman et al. 2008, Goes et al. 2014). To influence the motivation of posting negative reviews, identity disclosure may be essential because the mechanism works through reviewers' aversion of social disapproval (Huang et al. 2017). However, the review sites in our context have limited identity disclosure information, in contrast to the rich identity information from contexts in the literature (Proserpio and Zervas 2017). Following this literature, the motivation argument predicts that MRs would increase review volume, but the impact on review valence may depend on the richness of social identity in the context.

An alternative mechanism is that MRs may affect future reviews through its impact on sales (Ye et al. 2008). Because MRs may highlight positive reviews and mitigate the impact of negative reviews, they may increase future sales given that consumers rely on reviews to evaluate products before purchase (Ho et al. 2017). Recent research shows that MRs may indeed increase sales (Kumar et al. 2018, Xie et al. 2014). This mechanism is related to the service recovery literature (e.g., Lewis and McCann 2004, Maxham III and Netemeyer 2002). In offline settings, service recovery focuses on complaining customers to prevent the negative effect from spreading (Kelley and Davis 1994). In an online setting, Gu and Ye (2014) show MRs positively influence *repeat* customers, though they also find that such responses decrease the satisfaction of *other* customers. However, since they focus on repeat customers, the impact of MRs on a wider audience, which is our focus here, is not clear. Since the previously mentioned explanations focus on MR's mitigating effect on negative reviews, we call this mechanism the *mitigation* argument.

We also add to the literature by showing that responding target and style are much more nuanced. Lappas et al. (2016) find that a large portion of reviews

do not receive responses, especially for negative ones. Even for those with feedback, few responses provide detailed justification. Xie et al. (2017) find that different roles of responders may matter for different classes of hotels. Merely responding is not enough—improper responses may even backfire. Instead, the manager needs to respond to the right type of reviews with the right style. For example, we find that even though MRs to positive reviews may increase future review volume, dwelling too much on the incident may decrease its positive effect. This result is consistent with the literature that more information may be worse on certain occasions (Shulman et al. 2015). We also find that when responding to negative reviews, managers need to be specific by either explaining what has happened with the customer or disclosing the improvements after the incident. Such differential effects of MRs on positive and negative reviews further support our *mitigation* argument. These results also provide actionable managerial implications for managers when they respond to online reviews.

3. Empirical Settings

3.1. Research Context

Our major empirical setting is Ctrip.com (Ctrip; NASDAQ: CTRP) and eLong.com (eLong, NASDAQ: LONG), which are the two largest online travel

agencies in China (iResearch 2014). Similar to Expedia in the United States, they offer many services, including flight tickets, hotel booking, and travel guides. Our study focuses on their hotel booking services, and in particular, the online reviews and MRs on their platforms.

Both travel websites allow customers to provide online reviews for hotel stays. Also, Ctrip allows the hotel management to give responses to customer reviews, whereas, during our study period, eLong does not. To show a concrete example, Figures 1 and 2 provide screenshots with translations on both websites of Sofitel Wanda Chengdu, a five-star hotel in Chengdu, China. They provide examples of online customer reviews, including ratings and review texts, as well as the MRs on Ctrip.

Both Ctrip and eLong spend significant efforts to ensure the quality and accuracy of online customer reviews. They allow only real customers to post reviews to the host hotel after each stay (one review per online transaction) in our study period. To encourage submission of online reviews, the travel agents send email reminders with review links to each customer's registered email address right after check-out. They also have dedicated teams to identify and remove questionable reviews. These quality control processes make it difficult for hotels to bias the information by posting

Figure 1. (Color online) A Sample Hotel Review Page on Ctrip with Managerial Responses

The screenshot displays the Ctrip.com website interface for the 'Sofitel Wanda Chengdu' hotel. At the top, the Ctrip logo and navigation links are visible. The main content area shows a list of reviews for the hotel. Each review includes a user profile, a rating (represented by a star icon), and the review text in both Chinese and English. A managerial response is also present, addressing the concerns raised in the reviews. A diagram on the left side of the screenshot identifies the 'Customer Reviews' and 'Managerial Response' sections.

Customer Reviews	Managerial Response
<p>Jesse2005 (More Reviews) Hotel: Sofitel Wanda Chengdu April-27-2008</p> <p>Rating: 5.0</p> <p>相当不错的服务, 房间也不错, 值得住。而且位置很不错。 I was very satisfied with the service. The room was also quite nice. It is worth it. The location is also very good.</p>	<p>酒店反馈/ Hotel Feedback April-11-2008:</p> <p>非常感谢您的光临和对我们饭店的关注和肯定。对于您提出的不足我们将努力改进, 力求把工作做得更好, 给宾客提供更优质、更舒适的硬件设施和软件服务。 Thanks for your stay with our hotel and your attention and positive comments. We will do our best to address your concerns. We strive to provide the most comfortable facilities and the best services to all our guests.</p>
<p>200171**** (More Reviews) Hotel: Sofitel Wanda Chengdu April-8-2008</p> <p>Rating: 5.0</p> <p>房间有法式风格, 尤其浴室和化妆间的布置。不过淋浴房的防水做的不好, 洗完澡浴室地面都是水, 差点摔跤。 The room has a French style, especially with the layout of bathroom and dressing room. However, the shower is not waterproof. I almost tumbled over water on the bathroom floor.</p>	
<p>David**** (More Reviews) Hotel: Sofitel Wanda Chengdu April-15-2008</p> <p>Rating: 5.0</p> <p>五星级酒店肯定是各方面都很好的, 下次还要预定这里。 All aspects of this five-star hotel are surely excellent. I will book this hotel again next time.</p>	
<p>pi**** (More Reviews) Hotel: Sofitel Wanda Chengdu March-23-2008</p> <p>Rating: 4.0</p> <p>可以, 不错, 好呀, ——就是有点贵呀 The hotel is good, but it is a bit expensive.</p>	

Figure 2. (Color online) A Sample Hotel Review Page on eLong

The screenshot displays the eLong.com website interface for the 'Sofitel Wanda Chengdu' hotel. At the top, the eLong logo and navigation menu are visible. The hotel's name is prominently displayed in both Chinese and English. Below the name, the hotel's address and business district are listed. A section titled 'Genuine Customer Reviews (25)' is highlighted, showing a list of reviews. Each review entry includes a recommendation status (e.g., 'Recommended' or 'Not Recommended'), a date, and a translated English summary of the review content. The reviews are sorted by date, with the most recent at the top.

promotional or fake reviews. Any hotels that intend to post fake reviews will have to book through the travel agencies and pay a sizable commission. In addition, given the substantial number of reviews on the travel agency sites, fake reviews (even if they slip through the process) are likely to have a very limited impact.

3.2. Identification Strategy

Our focus in this paper is to identify the effect of MRs on subsequent reviews, both in volume and valence. But this task is not trivial. A cross-section comparison between hotels with and without MRs may bias our estimates because the two groups may have different underlying qualities. A longitudinal comparison before and after the first MR may not accurately quantify the influence of MRs either because the provision of MRs could correlate with the growth of hotels, which could be related to higher review volume and valence. Even a difference-in-differences specification

on one site may not satisfactorily address the concern because the provision of MRs could be due to changes in underlying management expertise, such as the introduction of a new management team that is more willing to listen to customers and simultaneously improves responses and service quality.

To accurately assess the influence of MRs, it is necessary to control the time-varying managerial expertise that is not observable to researchers. Even though we do not observe such offline properties of hotels, a unique feature of this study is that we leverage the *same* hotels on the two travel agencies to isolate the issue. Since we observe the same hotel across the two sites, we can ensure that the unobserved hotel quality is controlled by taking the difference between the two platforms.

To measure the effect accurately, we rely on a DDD specification to identify the relationship. For the ease of discussion, we call hotels that have posted MRs “MR hotels,” and those that have never posted a managerial

response “non-MR hotels.” In our DDD specification, the first difference is between the measures from the same hotel on Ctrip versus eLong, which controls the unobserved hotel quality change. The second difference is between MR and non-MR hotels, which account for their inherent quality difference. In this process, the non-MR hotels serve as a control group. The treatment is the event that a hotel posts the first MR. Only MR hotels receive this treatment. The third difference is before versus after the treatment within MR hotels. The DDD design utilizes the difference between MR and non-MR hotels across Ctrip and eLong to control for the unobserved trend of underlying hotel quality.

3.3. Empirical Models

To control for other variables’ impact on subsequent customer reviews, we employ a regression DDD approach. The regression framework allows us to control for time-specific and hotel-specific effects in our panel data. Our specification is as follows:

$$\Delta Y_{it} = \gamma MR_i + \beta MR_i \times After_{it} + \delta' X_{it} + \alpha_i + \theta_t + \varepsilon_{it}, \quad (1)$$

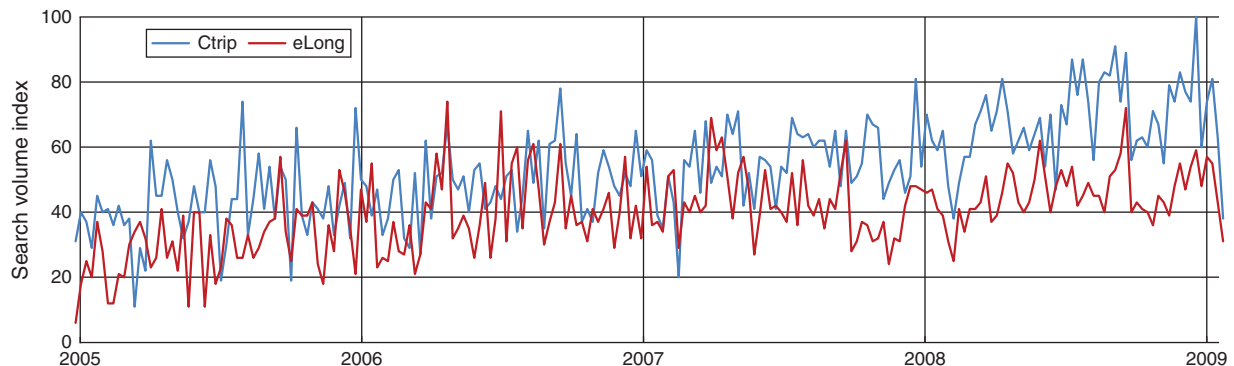
where ΔY_{it} is the difference of review volume (or valence) across Ctrip and eLong for hotel i at month t , MR_i identifies whether hotel i has provided MR by the end of our study period, $After_{it}$ captures whether hotel i has started to provide managerial response by month t , and X_{it} is a vector of control variables. Note that some hotels post MRs a long time after the review has been published. For these hotels, we use the time that the first MR been posted to identify $After_{it}$. In Equation (1), the coefficient γ captures the difference between MR and non-MR hotels. The coefficient β is our regression DDD estimator, which captures the effect of MR on the subsequent review valence and volume. We also include monthly dummies to control for the intertemporal dynamics in online review generation. We adopt robust clustered error terms at the hotel level to account for autocorrelation in the data across hotels and over time (Bertrand et al. 2004).

There are several potential concerns with the specification in Equation (1). First, each hotel may have its characteristics that correlate with the decision whether or not to provide MR. Thus, we also adopt a fixed-effect (FE) model to account for the unobserved hotel-specific effect. Note that in the FE model MR_i drops out because it is hotel specific and does not vary over time.

Second, the two agencies may have different underlying trends that render even the same hotel on the two platforms incomparable. The concern is twofold. One issue is that whether the two platforms have different designs that might affect the reviews. For example, if Ctrip only allows reviews from real customers and eLong allows reviews without staying, then the nature of reviews may be different. According to Figures 1 and 2, both platforms only allow reviews from customers who have booked through them. Hence, third-party reviews may not be a concern in our sample. The other issue is that the two agencies may serve different customer segments or have different trends. We compare their popularity based on Google Trends in Figure 3. The figure shows that the two sites attract a similar level of consumer interests, but Ctrip became increasingly popular in later years. To address the concern on the changing popularity of the two sites, we add month dummies into the vector X_{it} to account for the potentially different trends. If the different trends happen for all hotels across Ctrip and eLong, then this time trend variable controls for the difference.

The third concern is that non-MR and MR hotels may have different trends regarding their differences across Ctrip and eLong. For example, if customers of MR hotels were migrating from eLong to Ctrip more than those from non-MR hotels, the review volume may increase more for MR hotels. This migration may also correlate with the provision of MRs. To test such possible different time trends between non-MR and MR hotels, we add a time trend variable $Month_t$ and an interaction term $MR_i \times Month_t$ in the control vector X_{it} to test it. The coefficient of the interaction term will capture the different time trends between the two groups.

Figure 3. (Color online) Google Trends for Ctrip and eLong



Finally, hotels may have started to post MRs because of some random shock of negative ratings. That is, the appearance of MR may come from the negative ratings of MR hotels comparing to non-MR hotels. In this case, the ratings may return to the mean (increase) even if nothing happens. In the literature, this pretreatment trend is called the Ashenfelter's dip (Ashenfelter and Card 1985), which is common in difference-in-differences specifications. To address this concern, we first adopt the relative time model (Angrist and Pischke 2008) to examine whether there exists an Ashenfelter's dip in our sample. We also confirm our results by excluding immediate observations before and after the first MR.

4. Data

4.1. Data Collection

The data in this study were retrieved from the websites of Ctrip and eLong. We developed two Java crawlers to retrieve all available information about hotels on Ctrip and eLong from 10 major cities in China.² For each hotel in our data set, we collected the review contents, review ratings, posting dates, the presence (or lack) of MR, and the MR content to each customer review. We collected data for the period between January 2006 and July 2011. But since eLong started to provide the managerial response function at the beginning of 2009 and Ctrip underwent several design changes in 2009, we only use the data from January 2006 to December 2008 to ensure that our identification strategy is valid. In total, our sample contains 2,733 hotels with 72,645 reviews from Ctrip, and 1,759 hotels with 11,570 reviews from eLong. We match the hotels on Ctrip and eLong using hotel names and phone numbers, and check the matched list manually to ensure accuracy. The combined matching approach identifies 943 hotels that are available on both sites in our sample.

We then aggregate the review data to the hotel-month (i, t) level for our main analysis. For each hotel i in month t , we calculate the total number of reviews

and mean rating on both agencies. Note that some hotels may have no reviews on Ctrip or eLong in certain months. In our aggregation, setting the valence to zero may negatively bias the rating. Hence, we replace the missing mean valence with the historical mean valence for hotel i up to month t .³

We also note that eLong has a different rating system from Ctrip (binary versus five-point ratings) as shown in Figures 1 and 2. This difference may affect our results on the influence of MRs on valence in the matched hotel estimation. Moreover, some reviews on eLong have only review text but no ratings (as the third review shown in Figure 2). To address this concern, we use a text mining approach to extract sentiments of each review. We use a text mining classifier, *Dynamic Language Model* classifier with n -grams ($n = 16$), from the LingPipe suite to perform the estimation.⁴

4.2. Summary Statistics

Table 1 presents the summary statistics of the 943 hotels in our matched sample. The variables Vol_{it} and Val_{it} capture review volume and mean valence from Ctrip, respectively. $eVol_{it}$ and $eVal_{it}$ represent the volume and valence of reviews on eLong. We observe that on average Ctrip has higher volume of reviews than eLong. Meanwhile, the valence of reviews is similar on both agencies, with $eVal_{it}$ slightly higher. We also calculate the historical mean valence ($CumVal$) and cumulative volume ($CumVol$) of reviews for each hotel on each site to control for the influence of past reviews.

Mean value of MR_i shows that about 21% hotels have provided MRs. In our sample, 197 out of 943 matched hotels have adopted MR by the end of our sample period. Since $MR_i \times After_{it}$ is the interaction of the two variables, we can calculate that MR hotels on average have MRs for about 28% ($=0.06/0.21$) of the sample periods. After the adoption, 1,169 out of 4,133 new reviews receive MRs. This proportion means that even among MR hotels, there exist large variations in their MR practice. We later use the heterogeneity of MRs to conduct additional analyses.

Table 1. Summary Statistics of Matched Hotels Before Differencing

Variable	Description	Mean	sd	Min	Max
Vol_{it}	Number of new reviews on Ctrip	1.33	2.17	0.00	36.00
Val_{it}	Mean valence of reviews on Ctrip	0.29	0.32	0.00	1.00
$CumVol_{it}$	Total number of reviews on Ctrip	35.54	36.39	0.00	282.00
$CumVal_{it}$	Cumulative mean valence on Ctrip	0.17	0.13	0.00	1.00
$eVol_{it}$	Number of new reviews on eLong	0.36	0.87	0.00	25.00
$eVal_{it}$	Mean valence of reviews on eLong	0.37	0.35	0.00	1.00
$eCumVol_{it}$	Total number of reviews on eLong	9.38	8.81	0.00	80.00
$eCumVal_{it}$	Cumulative mean valence on eLong	0.37	0.28	0.00	1.00
MR_i	Whether hotel i has adopted MR in our sample	0.21	0.41	0.00	1.00
$MR_i \times After_{it}$	Whether hotel i has adopted MR by month t	0.06	0.24	0.00	1.00

Note. Subscripts index for hotel i and month t .

Table 2. Comparison of Means Between Non-MR and MR Hotels

Variable	(1) Non-MR hotels	(2) MR hotels— Before	(3) MR hotels— After	(4) <i>t</i> -stats	(5) <i>p</i> -value
ΔVal_{it}	−0.075	−0.095	−0.080	1.046	=0.30
ΔVol_{it}	0.864	1.187	1.903	8.791	<0.01

5. Results

5.1. Model-Free Evidence

In this section, we examine the impact of MRs on the valence and volume of subsequent reviews. We first report the differences of volume and valence from the two agencies for MR and non-MR hotels in Table 2. We calculate the differences by $\Delta Vol_{it} = Vol_{it} - eVol_{it}$ and $\Delta Val_{it} = Val_{it} - eVal_{it}$. In column (1) to (3), the mean values for the differences are presented for non-MR hotels, MR hotels before adoption, and MR hotels after adoption, respectively. We find that the valence difference of MR hotels is similar to that of non-MR hotels both before and after the first MR. Regarding review volumes, MR hotels have a higher number of reviews on Ctrip before their first MR than non-MR hotels, and they further experience a large increase in review volume after adopting MR. We conduct two-sided *t*-test for columns (2) and (3) and report the *t* statistics and *p*-value in columns (4) and (5). Both tests have the null hypotheses that MR hotels have the same value for the variable before and after adoption of MR. It is rejected that the volume differences are the same before and after the first MR, while we cannot reject the hypothesis that the valence differences are the same.

5.2. Impact on Review Valence and Volume

We next run our regression DDD specification in Equation (1) and report the results in Table 3. To reduce skewness of data and facilitate easy interpretation, we log-transformed the volume and valence of reviews

from the two agencies before we take the difference.⁵ The first three columns show the impact of MR on review volume. Model (1) reports the results from ordinary least squares (OLS) regression, with robust standard errors clustered at the hotel level. According to the coefficient of $MR_i \times After_{it}$, hotels on average have 14.2% ($P < 0.01$) more reviews per month after the adoption of MR. The potential different time trends of Ctrip and eLong is captured by the month dummies. Because the previous volume and valence may drive the sales and subsequent reviews, we control the cumulative volume and valence on each site following the literature (Chevalier and Mayzlin 2006, Chintagunta et al. 2010). These two variables enter Table 3 as $\Delta CumVol_{i,t-1}$ and $\Delta CumVal_{i,t-1}$ after the difference.

In model (2), we include linear time trend variables for our treatment and control groups to test whether MR and non-MR hotels have different trends. The potential different time trends of Ctrip and eLong are controlled by $Month_t$. The insignificant and small coefficient of $MR_i \times Month_t$ indicates that there is otherwise no significant difference over time between our treatment and control groups.⁶ Model (3) shows results with our FE model. The coefficients of the DDD estimator $MR_i \times After_{it}$ is highly similar to that in model (1).

Columns (4)–(6) of Table 3 report the impact of MR on review valence. Models (4) and (6) report results from the OLS regression and FE model, respectively. We do not find a significant valence increase after a hotel starts to provide MRs. Model (5) includes the linear time trend term $Month_t$ and interaction term $MR_i \times Month_t$ to test the different trends of MR and non-MR hotels. Similar to model (2), we cannot reject the hypothesis that the two groups have the same time trend. To save space, we only report FE model results for the rest of the paper.

5.2.1. Relative Time Model and Ashenfelter's Dip. In this section, we report our formal analysis of the treatment effect before and after the treatment (adoption

Table 3. Impact of MR on Subsequent Reviews

DV	$\Delta \log Vol_{it}$			$\Delta \log Val_{it}$		
	(1)	(2)	(3)	(4)	(5)	(6)
MR_i	0.069***	0.102**		0.030	0.070	
$MR_i \times After_{it}$	0.142***	0.147***	0.123***	0.071	0.072	0.092
$\Delta CumVol_{i,t-1}$	0.005***	0.005***	−0.008***	0.003***	0.003***	−0.004***
$\Delta CumVal_{i,t-1}$	−0.030	0.007	0.145***	1.882***	1.893***	1.462***
$Month_t$		−0.007***			−0.002	
$MR_i \times Month_t$		−0.002			−0.002	
Hotel dummies	No	No	Yes	No	No	Yes
Month dummies	Yes	No	Yes	Yes	No	Yes
Observations	23,082	23,082	23,082	23,082	23,082	23,082
Adjusted R^2	0.169	0.066	0.147	0.138	0.135	0.029
Model	OLS	OLS	FE	OLS	OLS	FE

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

of MR in our case). As previously mentioned, a major concern is the Ashenfelter's dip, which states that the pretreatment trend of the treatment group may bias the estimates in a difference-in-differences setting (Ashenfelter and Card 1985). Following the common practice in difference-in-differences studies, we adopt the relative time model (Angrist and Pischke 2008), which is used to check if a pretreatment trend exists for hotels. In our context, the model is specified as

$$\Delta Y_{it} = \gamma MR_i + \sum_{\tau=-T}^T \lambda_{\tau} MR_i \times D_{i\tau} + \delta' X_{it} + \alpha_i + \theta_t + \varepsilon_{it}, \quad (2)$$

where $D_{i\tau}$ is a dummy variable indicating whether the current month t is τ month before (in the case of a negative τ) or after (in the case of a positive τ) the month hotel i adopts MR. Equation (2) is similar to Equation (1), just with $MR_i \times After_{it}$ replaced by a set of dummy variables $MR_i \times D_{i\tau}$ for MR hotels. The set of coefficients λ_{τ} can help us identify whether there exists a pretreatment trend and how the effect changes after the treatment.

We estimate Equation (2) with $T = 9$ and visualize the evolution of the treatment effect in Figure 4, with panel (a) for volume and panel (b) for valence.⁷ In both panels, the X axis is the months relative to MR adoption, with the adoption month equal to 0. The solid line connects the estimated coefficients λ_{τ} ($\tau = -9, -8, \dots, -1, 0, 1, \dots, 9$), and the dashed lines connect the 95% confidence intervals of these coefficients. We can use the value of the coefficients and the confidence intervals to evaluate whether there exists a pretreatment trend in the volume and valence.

We observe two features of the review dynamics before MR adoption. First, review volume trends for MR and non-MR seem to be parallel before the treatment, as shown in panel (a). The only exception is that there seems to be a significant increase of review volume a few months before hotels adopt MR, though only the month before MR adoption ($\tau = -1$) have a statistically significant increase (the lower bound of the 95% confidence interval is greater than zero). Second, the review valence trends seem to be parallel for the control and treatment groups before the treatment. We also observe a slight dip in valence in the month right before hotels adopt MR. These two observations suggest that hotel management starts to provide MR as a reaction to a sudden increase (about 20%) of negative reviews. After hotels start to provide MR, the volume increases between 10% to 20% (with the lower bounds of 95% confidence interval greater than zero for most months after adoption of MR), which is consistent with our estimation in Table 3. Meanwhile, valence does not seem to be significantly affected.

Even though we find some evidence of Ashenfelter's dip, we note that the common concerns may not apply in our case. For review volume, our model may have underestimated the effect of MRs because the volume may have started to increase before the adoption of MR. Regarding review valence, if a hotel experiences a decrease in rating right before it starts to provide MR, the valence after MR may revert to the mean even if there is no MR. Hence, the presence of Ashenfelter's dip may overestimate the effect of MR on review valence. However, we do not find a significant increase in valence after MR, which means such concerns may not be a big problem in our case.

Figure 4. Evolution of Treatment Effect

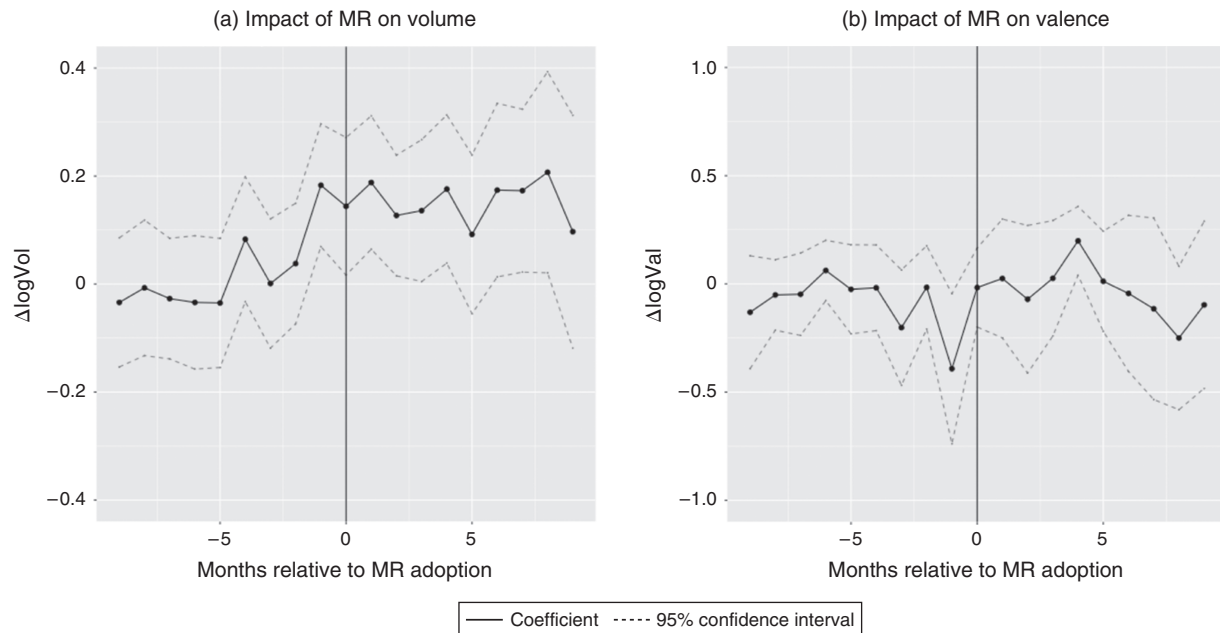


Table 4. Impact of MR After Correcting for Ashenfelter's Dip

DV	$\Delta \log Vol$			$\Delta \log Val$		
	(1)	(2)	(3)	(4)	(5)	(6)
$MR_i \times After_{it}$	0.141***	0.161**	0.140***	0.035	-0.025	-0.047
$\Delta CumVol_{i,t-1}$	-0.008***	-0.008***	-0.008***	-0.004***	-0.003***	-0.003***
$\Delta CumVal_{i,t-1}$	0.149***	0.151***	0.152***	1.468***	1.456***	1.467***
Month dummies	Included					
Observations	22,522	21,889	22,544	22,522	21,889	22,544
Adjusted R^2	0.147	0.146	0.147	0.031	0.031	0.032

Notes. In models (1) and (4), observations one month before and after MR adoption ($\tau = -1, 0$ and 1) are excluded. In models (2) and (5), observations three months before and after MR adoption are excluded ($\tau = -3, -2, -1, 0, 1, 2, 3$) are excluded. In models (3) and (6), observations three months before MR adoption are excluded ($\tau = -3, -2, -1$).

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

We nonetheless follow the common practice to correct for Ashenfelter's dip by excluding the sample periods around the adoption of MR. Specifically, we test the robustness of our results by excluding the observations one month and three months around MR adoption, as well as just three months before the adoption. By removing the pretreatment trends, the coefficients on review volume should increase while the ones on review valence should decrease. Our results after correcting for Ashenfelter's dip are reported in Table 4. For review volume, the coefficients for $MR_i \times After_{it}$ indeed increase in models (1)–(3) relative to the FE results in Table 3. The impact of MR on valence is still statistically insignificant across models (4)–(6) but decreases compared to the results in Table 3. All other results are consistent with results in Table 3.

5.2.2. Alternative Explanations and Heterogeneous Treatment Effect. We next affirm that it is MRs that change the subsequent review volume. Even though we control for the underlying hotel quality by differencing across Ctrip and eLong, it is possible that the migration of customers to Ctrip rather than the presence of MRs has caused the change in review volume. These two explanations have different predictions after the first MR. If the change comes from the migration of customers to Ctrip, then the presence and frequency of subsequent MRs should not matter. Otherwise, the presence of MRs should influence subsequent reviews because we do not expect one MR to have a lasting effect on all subsequent reviews.

We test these two alternative explanations by exploiting the display mechanism of reviews on Ctrip. The reviews on Ctrip are ordered by time and displayed in pages. Therefore, older reviews move into deeper pages when new reviews come in. We argue that the effect of MR may only exist when it is presented on the first several pages because few people read beyond two pages (Pavlou and Dimoka 2006). Since MRs are displayed with their corresponding reviews, the effect

of MRs may disappear once they move past the first few pages.

To statistically test this, we construct a variable, $Since_MR_{it}$, which measures how many months have passed since the *last* MR of the hotel. For example, if a hotel responded to five reviews in August 2007 and stopped, then $Since_MR_{it}$ is 1 for September 2007, and 2 for October 2007. A new MR in November 2007 will reset it to 0. And it will increase again if there is no further MRs. We add $Since_MR_{it}$ to the vector X_{it} in our DDD regression specification in Equation (1), and reestimate the model. The results are shown in models (1) and (3) of Table 5. Model (1) shows that as time elapses from the last MR, the effect on review volume begins to fade. On average, the volume would decrease for about 1.8% ($P < 0.01$) per month, which will take 10 months for the positive impact of adopting MR (18.3%, $P < 0.01$) to disappear if no further MRs are provided. If we consider the number of reviews per month, this is roughly the time all MRs move to the third page. Meanwhile, this effect is not significant on review valence as depicted in model (3) of Table 5.

The effect of $Since_MR_{it}$ suggests that heterogeneity in MR practices may affect subsequent reviews. We further examine it with total number of MRs in models (2) and (4) in Table 5, where $TotalMR_i$ is the total

Table 5. Time and Heterogeneous Effects of MR

DV	$\Delta \log Vol$		$\Delta \log Val$	
	(1)	(2)	(3)	(4)
$MR_i \times After_{it}$	0.183***	0.055	0.033	0.100
$Since_MR_{it}$	-0.018***		0.018	
$MR_i \times After_{it} \times TotalMR_i$		0.010***		-0.001
$\Delta CumVol_{i,t-1}$	-0.008***	-0.008***	-0.004***	-0.004***
$\Delta CumVal_{i,t-1}$	0.154***	0.138***	1.453***	1.463***
Month dummies	Included			
Observations	23,082	23,082	23,082	23,082
Adjusted R -squared	0.148	0.149	0.030	0.029

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

number of responses that hotel i provides after MR adoption.⁸ As shown in model (2), we find that the marginal effect of a single MR on review volume is about 1%, while the overall effect of adoption is taken away by the number of MRs. We do not find a significant effect of total MR number on review valence. With the results in Table 5, we conclude that it is the presence of MR rather than customer migration that influences the subsequent review volume.

5.3. Mechanisms of Managerial Responses

From our previous analysis, we can seemly conclude that the adoption of MR increases subsequent review volume but has little effect on valence. However, it is still not clear why MRs have such effects. In this section, we investigate the mechanism of MR by examining the heterogeneous MR practices among hotels after they adopt MR. We first examine whether MR has different effects on positive and negative reviews. We then investigate whether the response target and style matter.

5.3.1. Impact on Positive and Negative Reviews. We first examine the impact of MR adoption on the number of positive and negative reviews, respectively. As we discussed in the literature review, there may be two mechanisms through which MR affects reviews. With the *motivation* argument, MRs should increase the overall volume of reviews, and potentially decrease negative reviews if social identity is presented. With the *mitigation* argument, we should see an increase in both positive and negative reviews because MRs work through increased sales and should not change the relative composition of positive and negative reviews.

To test our hypotheses about these two mechanisms, we construct three new dependent variables and run our regression DDD specification in Equation (1), respectively. The results are shown in Table 6, where $\Delta \log PosVol$ is the difference of log-transformed volume of positive reviews between Ctrip and eLong, $\Delta \log NegVol$ is calculated for negative reviews in a similar fashion, and $\Delta \log PosRatio$ computes the difference of ratios of positive reviews over all reviews. Note that the positive and negative reviews are identified through our text mining process because of the different settings and scales on the two sites.

Table 6. Impact of MR on Positive and Negative Reviews

	(1) $\Delta \log PosVol$	(2) $\Delta \log NegVol$	(3) $\Delta \log PosRatio$
$MR_i \times After_{it}$	0.116***	0.099***	0.012
$\Delta CumVol_{i,t-1}$	-0.003***	-0.004***	-0.002***
$\Delta CumVal_{i,t-1}$	-0.049*	0.062**	-0.003
Month dummies	Yes	Yes	Yes
Observations	20,839	21,963	21,543
Adjusted R-squared	0.149	0.086	0.035

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

According to models (1) and (2) in Table 6, we find that the volumes of both positive and negative reviews increase on Ctrip compared to the corresponding numbers on eLong. The number of positive reviews increases by about 11.6%, while the number of negative reviews is 9.9% higher. To test whether the proportion of positive reviews increases relative to negative reviews, we use the ratio of positive reviews as the dependent variable in model (3). We do not find a statistically significant increase in positive review percentage. Therefore, the impact of MR seems to increase both positive and negative reviews at the same time, supporting the *mitigation* argument. However, we cannot completely reject the *motivation* argument because of a context difference in our setting. As shown in Figures 1 and 2, neither Ctrip or eLong disclose identity information of reviewers on the review page (the user IDs are anonymized with asterisks). This design difference can be considered as a different treatment in the review environment in contrast to the literature (Proserpio and Zervas 2017) where identity is revealed. The identity disclosure may be crucial to *discourage* negative reviews (Forman et al. 2008, Huang et al. 2017, Proserpio and Zervas 2017). Hence, in our context, the lack of identity disclosure works in favor of customers who post negative reviews (thus more truthfully reveal review sentiment). MRs increase subsequent review volume either through general motivation of posting reviews, or through increased sales by mitigating the impact of negative reviews.

5.3.2. Target and Style of Managerial Responses. Insofar we have treated MR as homogenous and counted only the numbers, but the MRs provided by hotels may exhibit extensive heterogeneity. For example, hotels may selectively reply to certain types of reviews, and they may write brief or detailed comments when they do respond. In this section, we further examine the heterogeneity by constructing several variables about the target and content of MRs. The summary statistics of these variables are presented in Table 7.

Table 7. Summary Statistics of MR Target and Style

Variable	Description	Mean	sd	Min	Max
$PosMR_{it}$	Number of MRs to positive reviews	0.57	1.45	0	13
$NegMR_{it}$	Number of MRs to negative reviews	0.19	0.62	0	6
$PosRatio_{it}$	Proportion of MRs to positive reviews	0.19	0.38	0	1
$NegRatio_{it}$	Proportion of MRs to negative reviews	0.12	0.32	0	1
$LenMR_{it}$	Log-transformed average length of MRs	1.15	1.89	0	6.38

Note. Summary statistics are calculated from MR hotels after they adopt MR.

Table 8. Impact of MR Target and Style

DV	$\Delta \log Vol$			$\Delta \log Val$		
	(1)	(2)	(3)	(4)	(5)	(6)
$MR_i \times After_{it}$	0.095**	0.094**	0.096**	0.059	0.060	0.060
$PosMR_{i,t-1}$	0.071**	0.368***		0.020	0.104	
$NegMR_{i,t-1}$	-0.007	-0.528**		-0.098	-0.403	
$LenMR_{i,t-1}$	0.004	0.004	0.027	0.039	0.035	0.040
$PosMR \times LenMR_{i,t-1}$		-0.074***			-0.020	
$NegMR \times LenMR_{i,t-1}$		0.127**			0.075	
$PosRatio_{i,t-1}$			0.700***			0.200
$NegRatio_{i,t-1}$			-0.517*			-0.606
$PosRatio \times LenMR_{i,t-1}$			-0.176***			-0.020
$NegRatio \times LenMR_{i,t-1}$			0.129*			0.077
$\Delta CumVol_{i,t-1}$	-0.008***	-0.008***	-0.008***	-0.004***	-0.004***	-0.004***
$\Delta CumVal_{i,t-1}$	0.143***	0.143***	0.145***	1.462***	1.462***	1.462***
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,082	23,082	23,082	23,082	23,082	23,082
Adjusted R-squared	0.148	0.148	0.148	0.030	0.029	0.030

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

In Table 7, $PosMR_{it}$ and $NegMR_{it}$ represent the number of MRs to positive (four–five star) and negative (one–three star) reviews for hotel i in month t , respectively. Note that these two variables only have positive values for the 197 MR hotels after they adopt MR. Hence, the summary statistics are calculated on this subsample. We can see that on average, hotels have more responses to positive reviews than to negative reviews. However, this may be because of the number of available positive and negative reviews. We then calculate the ratio of MRs to positive and negative reviews in $PosRatio_{it}$ and $NegRatio_{it}$, correspondingly.⁹ We can see that the ratios are more balanced, even though the ratio of response to positive reviews is still higher. We examine the style by counting the characters in each MR. The average length is then calculated and log-transformed for each hotel i in month t as the variable $LenMR_{it}$, which measures how detailed the responses are. We next use the MR target and style variables to examine the possible mechanisms. We follow similar specifications to Equation (1) but add lagged MR style variables. The results are presented in Table 8. Similar to our major results in Table 3, models (1)–(3) use review volume as the dependent variable, and models (4)–(6) use review valence as the dependent variable. Because most coefficients for valence are not statistically significant, our discussion focuses on models (1)–(3). In model (1), we add the lagged number of MRs to positive and negative reviews, as well as the average length of MRs. The results show that responding to positive reviews has a significant ($P < 0.05$) positive effect on the subsequent review volume. The coefficients for $NegMR$ and $LenMR$ are not statistically significant. The main effect of MR adoption is still positive, even though slightly decreases in magnitude.¹⁰

We further examine the effect of MR style on responses to different types of reviews by interacting $LenMR$ with $PosMR$ and $NegMR$ in model (2). After adding the interaction terms, several interesting results emerge. First, the effect of $PosMR$ becomes larger and more significant, and the coefficient of $NegMR$ becomes negative and significant ($P < 0.05$). While responses to positive reviews may highlight the positive features of the hotel, MRs to negative reviews risk emphasizing the mistakes and errors. Second and more interestingly, the length of the response weakens the main effects for both variables. On the one hand, lengthy MRs to positive reviews may attenuate the positive effect of the response, probably because longer comments contain unnecessary explanations that would direct future customers' attention to the negative aspects, even though the overall evaluation is positive. On the other hand, more detailed responses to negative reviews may mitigate the negative impact of the reviews. The differential effects of responses to positive and negative reviews partially support the *mitigation* argument that MR works through sales because otherwise MRs to both positive and negative reviews should have a positive effect on the volume by increasing the motivation to post reviews.

In both cases, the signs of the main effect may change when the $LenMR_{it}$ surpasses certain thresholds. For $PosMR_{it}$, the sign will flip when $LenMR_{it}$ is around 5, which is about 148 Chinese characters for the original message. And for $NegMR_{it}$, the threshold is around 4 (about 55 characters). One concern of using the response number is that the results may be biased by the availability of positive or positive reviews. Hence, we replace the number with ratio in model (3). The results are qualitatively consistent with and similar in

magnitude to those in model (2), just with both thresholds now close to 4 (about 55 characters). This means that when providing feedback to positive reviews, a longer than threshold response may backfire. Meanwhile, a longer than threshold response to negative reviews may address the concerns and turn the effect to positive.

5.4. Robustness Checks

Two concerns with our results may be related to our treatments of review valence. First, to make the valence comparable across Ctrip and eLong, we have used text mining to obtain the sentimental orientation of each review. However, this process may introduce biases on valence because the text mining technique is not perfect. Therefore, we run the same specification on user ratings from Ctrip and eLong to double check our results. Note that Ctrip uses a five-star rating scheme while eLong adopts a binary system, we have to convert them to the same scale before we use them in our DDD specification in Equation (1). We choose to divide Ctrip's five-star ratings by five to convert them into values between 0 and 1.

We then employ user rating and its log-transformed version as the dependent variables, and present the results in Table A.3 in Appendix A in the online appendix. The effect of MR is insignificant in FE models (1) and (3). We also note that MR and non-MR hotels do not seem to have different time trends from models (2) and (4).

Second, our data aggregation process to the hotel-month level may create some concerns. As described in Section 4, when a hotel has no reviews in a given month, we replace missing values of review valence with the historical mean valence ($CumVal_{it}$ and $eCumVal_{it}$ in Table 2).¹¹ The issue with this approach is that the historical mean valence may smooth out the change of review valence after MR, which leads to the insignificant effect on valence. To address this concern, we run a difference-in-differences model on the *review level* where we do not need to deal with missing observation issues. The model is as follows:

$$Val_{ijt} = \gamma After_{ijt} + \lambda Ctrip_i + \beta Ctrip_i \times After_{ijt} + \delta' X_{ijt} + \varepsilon_{ijt}, \quad (3)$$

where Val_{ijt} measures the valence of review i for hotel j in month t , $After_{ijt}$ indicates whether review i at time t is after hotel j adopts MR, $Ctrip_i$ indicates whether review i is from Ctrip, X_{ijt} is a set of control variables. The coefficient γ captures the review difference before and after adoption of MR, the coefficient λ captures the systematic difference between Ctrip and eLong, and the coefficient β captures the additional change of review valence on Ctrip after the adoption of MR. β is our difference-in-differences estimator that identifies the effect of MR on review valence.

We estimate Equation (3) with the review valence from our text mining process, with both the original sentimental orientation and the log-transformed value. The results are presented in Table A.4 in the online appendix. Models (1) and (3) estimate Equation (3) with FE models and cluster the robust standard errors on the hotel level. And models (2) and (4) additionally control for platform-specific time trends. All models include month dummies to control for the common shock to both platforms in specific time periods. In all specifications, our primary estimator β is positive but not significant, which confirms our conclusion on the nonsignificant effect of MRs on review valence.

6. Discussion and Conclusions

6.1. Major Findings

Our results show that MRs increase the subsequent review *volume*. In fact, the volume increases for about 12% to 14% after the provider begins to offer MRs. We also find that the influence of MRs dies out if no subsequent MRs are provided, and higher response activities after the adoption are associated with higher review volume. This increase may be because reviewers have a higher *motivation* to post reviews when they anticipate their comments to be valued, or because MRs *mitigate* the impact of negative reviews on future sales. The former *motivation* argument supports the literature on drivers of user-generated content in online communities in general (e.g., Chen et al. 2018), and drivers of online reviews in particular (Goes et al. 2014, Huang et al. 2017). Our findings add to the literature by showing that MRs can be an effective way to increase review volume, which may further improve future sales. The later *mitigation* argument draws upon the recent literature regarding the impact of MRs on sales (Kumar et al. 2018, Xie et al. 2014), and explores further the potential mechanisms that MRs affect customer decisions.

Our further analysis provides several hints for the underlying mechanisms. First, in contrast to previous research, we do not find a significant effect of MRs on review *valence*. Previous studies argue that MRs increase the motivation to post positive reviews while decreasing the incentive to post negative ones. As a consequence, the review valence would increase, and reviewers would respond with fewer but longer negative reviews (Proserpio and Zervas 2017). However, we find that both positive and negative reviews increase after MR adoption in our setting. It can be attributed to the *mitigation* argument because the presence of MRs may only affect the purchase decision. The underlying service quality may have remained the same or been controlled in our setting, so that the review valence remains unchanged. But we cannot exclude the *motivation* argument where the motivation for posting both positive and negative reviews can increase, especially considering a contextual difference in our setting from

the previous studies. Our research context did not disclose reviewers' identity, while such identity information is rich in prior studies (especially Proserpio and Zervas 2017). It turns out that this design matters in its impact on review valence. This difference reveals an important design element that may change the effect of MRs and community engagement in general, and can be an interesting research direction in the future.

Second, we find that MRs to positive reviews have a positive effect on review volume, and this effect may be weakened or even canceled by lengthy responses. The positive effect of MRs may come from a stronger *motivation* for future positive review posters, or a highlight of the positive experience so that the future sales would increase. But the negative moderating effect of MR length seems to support the *mitigation* argument, because a detailed response should increase rather than decrease the motivation to post positive reviews. Instead, detailed MRs to positive reviews may incorrectly emphasize the negative points in already positive reviews, which may undermine the positive influence of responding to positive reviews.

Third, we find that MRs to negative reviews can have a negative effect on review volume, and this effect is attenuated or even reversed by detailed responses. These results also lean toward the *mitigation* argument. The negative effect may come from the risk of highlighting service failures that have caused the negative review. And the positive moderating effect of response length shows that a detailed sincere explanation of the incident and future corrections can *mitigate* the negative impact. Overall, it seems that the motivation argument does not paint a full picture of the influence of MRs, which may impact future reviews by highlighting the positive reviews and mitigate the negative reviews. Even though we cannot strictly disentangle the two because of the limitation of our data, these two mechanisms revealed in our analysis can provide interesting directions for future research.

6.2. Managerial Implications

Our results provide several managerial insights for the management of online reviews. First, MR may be an effective tool to stimulate future reviews. Compared to paying customers to write reviews or posting fake reviews, MR could be a more effective way to generate honest online reviews. It can potentially increase sales even though we could not distinguish it from motivation change in our study. But given the shown role of review volume on future sales in the literature, managers can influence future review and sales through their MR practices. Second, MRs should be a consistent practice instead of a one-time occurrence. Because older MRs will move to deeper pages as new reviews come in, it would be beneficial to keep several MRs on

the first page. Third, responding to positive and negative reviews may influence future review volume differently. Moreover, it might be helpful to be brief when responding to positive reviews, while it is important to explain the details of the situation or improvement for responses to negative reviews. Hence, the nuances in responding to positive versus negative reviews are quite important. Finally, MRs may have less impact on review valence when the reviewers do not have identity disclosure information. If the managers truly want to hear negative voices, it might be a good practice to keep the reviewers anonymous. Ultimately, improving the service quality with the feedback from customer reviews may be a more effective way to increase the review valence.

6.3. Limitations

Our analysis also has several limitations that are worth exploring in future research. First, while we have taken several steps to address possible selection biases, our model cannot control for all confounding factors. A fully randomized experiment will be useful to validate the findings. Second, our data set spans a period when hotels just began to pay attention to MRs. As the practice becomes more popular, the effect of MRs and the underlying mechanisms may have changed. Given the limit of our current data, we are unable to disentangle the changing mechanisms. It is a great topic open for future research. Third, although we argue that MRs may influence future review behavior through sales, our data on customer purchase decisions is limited. It would be valuable for future research to obtain sales data on a large scale to assess the influence of managerial response on future reviews through sales.

6.4. Closing Remarks

As online customer reviews increasingly affect consumer purchase decisions, managing online reviews has become an important task for businesses. In this research, we focus on identifying the relationship between managerial responses and subsequent reviews (volume and valence). Our analyses suggest that managerial responses can significantly influence subsequent customer review behavior. As such, managerial responses could be a valuable tool for businesses to interact with customers in the online world.

Although our research context is limited to online hotel reviews, the findings can be potentially generalized to other electronic commerce environments. Given the growth of third-party review intermediaries that enable consumers to distribute and exchange reviews on any business, product, or service, businesses have increasingly recognized the importance of taking proactive steps to manage online reviews. For example, Yelp.com, Newegg.com, and even Apple app store now allow businesses to respond to customer

reviews. This research shows the impact of managerial responses and reveals the mechanisms behind the practice, which may help businesses manage their responses.

We summarize the key contributions of this study as follows. First, our analysis measures the impact of MRs on review volume and valence and explores the mechanisms of MR. Prior literature on MRs and service recovery studies the impact of MRs on the complaining customers (e.g., Gu and Ye 2014, Lewis and McCann 2004). Our current study extends this literature into externality by showing that MRs can also impact the subsequent reviews of *other customers*. The mechanism is either through motivating customers to post reviews (Proserpio and Zervas 2017), or influencing their postings through sales (Kumar et al. 2018). We also add to the literature by showing that the influence on valence may not be evident if the review sites do not disclose the identity of reviewers. Our results highlight that in the age of social media, businesses should carefully interact with the customers because the interactions may have spillover effects on other customers and subsequently affect their purchase and review decisions. Such externality was much more limited in traditional offline settings.

Second, this study extends the literature on the drivers of customer reviews. Previous research has suggested that reviews may be driven by underlying sales (Dellarocas and Narayan 2006), reviewer experience (Ho et al. 2017), and motivation of posting (Goes et al. 2014, Huang et al. 2017, Lee et al. 2015). Our research adds MRs as another significant driver—our evidence indeed shows that MRs, especially consistent practice over time, tend to influence the volume of the reviews. Therefore, this paper suggests an alternative way for businesses to interact with customers and manage customer reviews through something they can control.

Last but not least, our results show the differential effects of MRs to positive and negative reviews, as well as the distinct moderating roles of MR style. The differential effects of MRs further reveal the potential mechanisms of MRs. Even though MRs may motivate future customers to post more reviews, they may be more likely to influence future reviews by emphasizing the positive and negative elements in previous reviews. The distinct moderating effects of MR length on MRs to positive and negative reviews provide specific guidance for businesses to provide MRs online: be brief while responding to positive reviews, and be specific while responding to negative ones.

In a broader sense, we also demonstrate the externality of online interactions with customers. Our results show that the interactions will not only affect the focal customers that the vendor is interacting with, but also influence the decisions of future customers at large

who will observe these interactions. Therefore, careful management of such interactions is essential for businesses to thrive in online environments. We hope the initial results in this paper will stimulate more research in this growing area.

Acknowledgments

The authors are grateful for the constructive comments and suggestions made by the senior editor, Kai-Lung Hui, the associate editor, Param Vir Singh, and three anonymous reviewers.

Endnotes

¹ See <http://www.amazon.com/gp/vine/help>.

² Shanghai, Beijing, Guangzhou, Shenzhen, Chongqing, Chengdu, Harbin, Xi'an, Lanzhou, and Lhasas.

³ This replacement of missing values may affect our conclusion on review valence. We double-check our analysis of review valence on the review level in our robustness checks to address this concern.

⁴ LingPipe suite is available at <http://alias-i.com/lingpipe>. We report the text mining process and results in Appendix B in the online appendix.

⁵ For example, variable Y is transformed with $\log(Y + 1)$. We also provide the results with the original measure in Table A.2 in Online Appendix A. The results are qualitatively consistent.

⁶ We have also checked the robustness by including both time trend and dummies. More time dummies are dropped because of the collinearity of the two, but results are consistent.

⁷ The full results are reported in Table A.1 in Online Appendix A. We also estimated the model with $T = 3, 6$, and 12 . The results are highly similar.

⁸ Note that we omit the lower order terms MR_i , $TotalMR_i$, $MR_i \times TotalMR_i$, $After_{it}$, and $TotalMR_i \times After_{it}$ in models (2) and (4). MR_i , $TotalMR_i$, and $MR_i \times TotalMR_i$ will be absorbed by the fixed effect. $After_{it}$ is equivalent to $MR_i \times After_{it}$ and $TotalMR_i \times After_{it}$ is equivalent to $MR_i \times After_{it} \times TotalMR_i$, and therefore they will be dropped in the regression.

⁹ When there are no reviews, we put these two variables as 0.

¹⁰ Note that because only MR hotels have positive values for these variables, the three variables are equivalent to $MR_i \times After_{it} \times PosMR_{i,t-1}$, $MR_i \times After_{it} \times NegMR_{i,t-1}$, and $MR_i \times After_{it} \times LenMR_{i,t-1}$. We omit the common $MR_i \times After_{it}$ for the ease of presentation and reading.

¹¹ We have also run the hotel level analysis by dropping missing observations. The results are consistent.

References

- Anderson M (2014) Local consumer review survey 2014 | Online reputation survey. *BrightLocal*. Retrieved May 7, 2015, <http://www.brightlocal.com/2014/07/01/local-consumer-review-survey-2014/>.
- Angrist JD, Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).
- Ashenfelter O, Card D (1985) Using the longitudinal structure of earnings to estimate the effect of training programs. *Rev. Econom. Statist.* 67(4):648–660.
- Berger J, Schwartz EM (2011) What drives immediate and ongoing word of mouth? *J. Marketing Res.* 48(5):869–880.
- Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1):249–275.

- Bickart B, Schindler RM (2001) Internet forums as influential sources of consumer information. *J. Interact. Marketing* 15(3):31–40.
- Burtch G, Hong Y, Bapna R, Griskevicius V (2018) Stimulating online reviews by combining financial incentives and social norms. *Management Sci.* 64(5):2065–2082.
- Chen W, Wei X, Zhu KX (2018) Engaging voluntary contributions in online communities: A hidden Markov model. *MIS Quart.* 42(1):83–100.
- Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.
- Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Sci.* 29(5):944–957.
- Dellarocas C, Narayan R (2006) A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth. *Statist. Sci.* 21(2):277–285.
- Duan W, Gu B, Whinston AB (2008a) The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *J. Retail.* 84(2):233–242.
- Duan W, Gu B, Whinston AB (2008b) Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems* 45(4):1007–1016.
- Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inform. Systems Res.* 19(3):291–313.
- Godes D, Silva JC (2012) Sequential and temporal dynamics of online opinion. *Marketing Sci.* 31(3):448–473.
- Goes PB, Lin M, Yeung C man A (2014) “Popularity effect” in user-generated content: Evidence from online product reviews. *Inform. Systems Res.* 25(2):222–238.
- Gu B, Ye Q (2014) First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production Oper. Management* 23(4):570–582.
- Ho YC(C), Wu J, Tan Y (2017) Disconfirmation effect on online rating behavior: A structural model. *Inform. Systems Res.* 28(3):626–642.
- Huang N(N), Hong Y, Burtch G (2017) Social network integration and user content generation: Evidence from natural experiments. *MIS Quart.* 41(4):1035–1058.
- iResearch (2014) *2014 China Online Travel Report* (iResearch Consulting Group, Shanghai, China).
- Kelley SW, Davis MA (1994) Antecedents to customer expectations for service recovery. *J. Acad. Marketing Sci.* 22(1):52–61.
- Kumar N, Qiu L, Kumar S (2018) Exit, voice, and response in digital platforms: An empirical investigation of online management response strategies. *Inform. Systems Res.* 29(4):849–870.
- Lappas T, Sabnis G, Valkanas G (2016) The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry. *Inform. Systems Res.* 27(4):940–961.
- Lee D, Hosanagar K, Nair H (2018) Advertising content and consumer engagement on social media: Evidence from Facebook. *Management Sci.* 64(11):5105–5131.
- Lee YJ, Hosanagar K, Tan Y (2015) Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Sci.* 61(9):2241–2258.
- Levy SE, Duan W, Boo S (2013) An analysis of one-star online reviews and responses in the Washington, D.C., lodging market. *Cornell Hospitality Quart.* 54(1):49–63.
- Lewis BR, McCann P (2004) Service failure and recovery: Evidence from the hotel industry. *Internat. J. Contemp. Hospitality Management* 16(1):6–17.
- Li X, Hitt LM (2008) Self-selection and information role of online product reviews. *Inform. Systems Res.* 19(4):456–474.
- Ma M, Agarwal R (2007) Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Inform. Systems Res.* 18(1):42–67.
- Maxham III JG, Netemeyer RG (2002) A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *J. Marketing* 66(4):57–71.
- Mayzlin D, Dover Y, Chevalier J (2014) Promotional reviews: An empirical investigation of online review manipulation. *Amer. Econom. Rev.* 104(8):2421–2455.
- Moe WW, Schweidel DA (2012) Online product opinions: Incidence, evaluation, and evolution. *Marketing Sci.* 31(3):372–386.
- Park SY, Allen JP (2013) Responding to online reviews problem solving and engagement in hotels. *Cornell Hospitality Quart.* 54(1):64–73.
- Pavlou PA, Dimoka A (2006) The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Inform. Systems Res.* 17(4):392–414.
- Proserpio D, Zervas G (2017) Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Sci.* 36(5):645–665.
- Ren J, Yeoh W, Shan Ee M, Popović A (2018) Online consumer reviews and sales: Examining the chicken-egg relationships. *J. Assoc. Inform. Sci. Tech.* 69(3):449–460.
- Shulman JD, Cunha M, Saint Clair JK (2015) Consumer uncertainty and purchase decision reversals: Theory and evidence. *Marketing Sci.* 34(4):590–605.
- Wang Y, Goes P, Wei Z, Zeng DD (2017) *Production of Online Word-of-Mouth: Peer Effects and the Moderation of User Characteristics* (Social Science Research Network, Rochester, NY).
- Wu F, Huberman BA (2008) How public opinion forms. Papadimitriou C, Zhang S, eds. *Internet and Network Economics*. Lecture Notes in Computer Science (Springer, Berlin), 334–341.
- Xie K, Kwok L, Wang W (2017) Monetizing managerial responses on TripAdvisor: Performance implications across hotel classes. *Cornell Hospitality Quart.* 58(3):240–252.
- Xie KL, Zhang Z, Zhang Z (2014) The business value of online consumer reviews and management response to hotel performance. *Internat. J. Hospitality Management* 43:1–12.
- Ye Q, Law R, Gu B (2009) The impact of online user reviews on hotel room sales. *Internat. J. Hospitality Management* 28(1):180–182.
- Ye Q, Gu B, Chen W, Law R (2008) Measuring the value of managerial responses to online reviews: A natural experiment of two online travel agencies. *ICIS 2008 Proc.*, 115.
- Ye Q, Law R, Gu B, Chen W (2011) The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Human Behav.* 27(2):634–639.
- Zhang C, Hahn J, De P (2013) Continued participation in online innovation communities: Does community response matter equally for everyone? *Inform. Systems Res.* 24(4):1112–1130.
- Zhu F, Zhang X(M) (2010) Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *J. Marketing* 74(2):133–148.