Marketing Science Institute Working Paper Series 2021

Report No. 21-108

# How Incumbents Beat Disruption? Evidence from Hotels' Responses to Home Sharing

Wei Chen, Karen Xie, Yong Liu

**How Incumbents Beat Disruption?**

**Evidence from Hotels' Responses to Home Sharing**

Wei Chen, Karen Xie, Yong Liu[1]

Data of Submission: December 2, 2019

*Abstract*

The disruption of sharing economy services to incumbent firms has attracted growing attention. Yet, the literature is silent on how incumbents respond to the rivalry and the resulted consequences. We investigate incumbent hotels' adjustment on quality after home sharing's entry using management responses, an online reputation marketing strategy to address feedback in customer reviews. Our method integrates quasi-experiments and machine learning to not only estimate hotels' response to home sharing's entry but also unveil the mechanism. We provide evidence on distinct responses to home sharing's entry across different hotel price segments, which lead to divergent performance outcomes in terms of customer satisfaction and sales. Hotels that respond more actively to customer reviews demonstrate improved quality in service areas where home sharing typically leads - including the check-in/out process, cleanliness, excursion opportunity, and room condition -and receive higher sales. In contrast, hotels that respond less appear to lose to not only home sharing but also peer hotels that respond more to reviews. We provide implications on how incumbents should react to technological and business model disruptions.

**Keywords***:* Sharing economy, Incumbent firms, Management responses, Difference-in-differences, Machine learning

---

[1] Assistant Professor of MIS at Eller College of Business, University of Arizona, weichen@arizona.edu; Associate Professor of Service Analytics & Betty and Fritz Knoebel Fellow at Daniels College of Business, University of Denver, Karen.Xie@du.edu; Marketing Department Head & Robert A. Eckert Endowed Chair in Marketing at Eller College of Business, University of Arizona, yoliu@arizona.edu. We thank the financial support from the Marketing Science Institute (www.msi.org) and the NET Institute (www.NETinst.org). For helpful comments, we are grateful to participants and/or reviewers at the Conference on Artificial Intelligence, Machine Learning, and Business Analytics 2019, the NET Institute Conference 2019, INFORMS Annual Conference 2019, Statistical Conference in E-Commerce Research (SCECR) 2019, Marketing Dynamics Conference 2019 and the seminar at University of Arizona and Hong Kong Polytechnic University. All results have been reviewed to ensure that no confidential information is disclosed. Three authors are randomly ranked. Any errors are ours.

## Statement of Intended Contribution

Amid the continuing disruption of sharing economy services to incumbent firms in the traditional hotel industry is the unaddressed curiosity of whether and how incumbent firms battle back. This paper fills the void by answering questions of particular interest to academics and practitioners: Whether and how do incumbent hotels respond to sharing economy's entry? Do incumbents' responses alter their performance outcomes? What is the mechanism through which incumbents beat disruptors? We focus on if and how the incumbent hotels adjust quality provision after home sharing' entry using management responses to customer reviews.

The variations in the timing of home sharing's entry across the vicinity of hotels represent an empirical opportunity to estimate the entry effect on hotels' management responses and performance outcomes (i.e., customer satisfaction and sales). Our estimation hinges on a quasi-experiment through a difference-in-differences design. Our data is rich with unstructured textual reviews of both hotels and their home-sharing counterparts. To unveil the performance-improving opportunities in reviews, we deploy state-of-the-arts machine learning algorithms integrating natural language processing for topic modeling via Latent Dirichlet Allocation and deep learning for sentiment analyses using Word2Vec and Convolutional Neural Networks.

Our analysis reveals distinct responses to customer reviews across different types of hotels facing home sharing's entry, which lead to divergent performance outcomes in terms of customer satisfaction and sales, and sheds light on the mechanisms behind these effects. We contribute to the literature on disruptive innovation, sharing economy, and management responses to customer reviews. We also generate new insights on several important issues, including how incumbents become adaptable in the face of disruption, the duo-effect of incumbents' reaction in not only defending external disruptors but also differentiating from internal incumbent peers, and how machine learning can be deployed at the business level to assist addressing these questions.

This research provides useful implications for practice. One of the marketing strategies advocated in this study is to learn from the ubiquitously available customer reviews when facing entries by disruptive innovation. We show how incumbents can combine management responses to exploit opportunities in customer feedback and sustain competitive advantage. Through responding to reviews and automating machine learning, different types of incumbents can extrapolate and scale the efficiency for focused, targeted, and actionable strategies specific to their needs to remain competitive in the market.

## Introduction

Sharing economy platforms (Airbnb, Uber, Lending Club, etc.) enable individuals to make earnings using their under-utilized resources to offer services that used to be provided by incumbent firms (hotels, taxis, banks, etc.). In particular, the rise of home-sharing services such as Airbnb has increasingly enticed customers out of hotel rooms to local residences (Farronato and Fradkin 2018; Li and Srinivasan 2019; Zervas, Proserpio, and Byers 2017) while boosting customer satisfaction (Zervas, Proserpio, and Byers 2015). The overwhelming success has made home sharing a travel industry juggernaut, pushing incumbent hotels to rethink the way they offer services.

Amid the disruption of sharing economy is, however, the unaddressed curiosity of whether and how incumbent firms battle back. The existing literature is silent on a few questions of particular interest to academics and practitioners: Whether and how do incumbents respond to sharing economy's entry? Do incumbents' responses alter their performance outcomes when facing the competition with sharing economy entrants? What is the mechanism through which incumbents beat disruptors? These questions motivate our study.

In this paper, we focus on hotels and investigate if and how they adjust quality provision in response to home sharing' entry using management responses. In the form of an open-ended piece of text, a management response is written by hotel managers and publicly displayed underneath consumer reviews (reviews, hereafter) that it aims to address. Literature has widely documented management responses as an online reputation marketing strategy for managers to address customer feedback (in the reviews) and identify opportunities to improve the service (Gu and Ye 2014; Wang and Chaudhry 2018). We are particularly interested in observing the change in management responses across incumbent hotels after home sharing's entry, how divergent response strategies differentiate hotels in performance, and what opportunities for quality adjustments that managers can learn by responding to reviews.

Our emphasis on management responses as a competitive tool for incumbents when facing competitive move is inspired by both theory and practice. Prior work on entry and incumbents has shown that prices generally fall in the face of competition because incumbents tend to rely on reduced prices to make sales and beat entrants (e.g., McCann and Vroom 2010; Seamans 2013; Simon 2005). We extrapolate from this finding and argue that incumbents would compete along dimensions other than price (e.g., quality). Because service quality in the hospitality industry is arguably the most important factor for gaining customers' confidence and competitive advantage (Ankur 2018), we conjecture that hotels have the incentive to compete by actively responding to and learning from reviews for performance-improving opportunities. In practice. hoteliers have reportedly lobbied local legislators to regulate home sharing[2] and, recently, copy-catted home sharing's offerings[3] in an attempt to compete, both with modest success. Advocating management responses as a competitive tool, we aim to bring evidence to inform the industry's defense. While home sharing's entry has typically seen as a *threat* to the profitability of incumbent hotels, we view home sharing's entry as an *opportunity*. We envision ubiquitous reviews as the most accessible, cost-effective source of inspirations for hotels to develop defense strategies, right from their backyard.

We collected large-scale, multidimensional data on hotels and home-sharing properties in a highly popular tourist market, Beijing, China, over a period from 2015 to 2017. During this period, Beijing experienced exponential growth of home sharing, which enables us to observe

[2] Source: https://www.nytimes.com/2017/04/16/technology/inside-the-hotel-industrys-plan-to-combat-airbnb.html
[3] Source: https://www.curbed.com/2017/10/12/16466882/hotel-airbnb-hyatt-oasis-collection-hospitality

the evolution of competition between hotels and home sharing. For each hotel, we obtained data on purchase-verified reviews, management responses to the reviews, as well as property characteristics (price, geographic coordinates, etc.). In the vicinity of each hotel property, we identify home sharing's entry using its first time-stamped, purchase-verified review and calculate its supply accordingly. We combine the data from multiple sources and construct a panel at the hotel by year-month level. Our final sample includes 6,678 hotels and 3,584 home-sharing properties.

The variations in the timing of home sharing's entry across the vicinity of hotels represent an empirical opportunity to estimate the entry effect. Our estimation of the home sharing's entry impact on hotels hinges on a quasi-experiment through a difference-in-differences (DID, hereafter) design, in which the hotels experiencing home sharing's entry is the affected group, whereas hotels that have not yet experienced home sharing's entry are the control group. Because our data set is rich in unstructured textual reviews on both hotels and their home-sharing counterparts, we are able to unveil hotels' quality gaps observed by customers in comparison with home sharing. We deploy state-of-the-arts machine learning algorithms integrating natural language processing for topic modeling via Latent Dirichlet Allocation (LDA, hereafter), and deep learning for sentiment analyses via Word2Vec and Convolutional Neural Networks (CNN, hereafter). In addition to these main analyses, we rule out alternative explanations to validate our DID results. We have also furnished the Web Appendixes to check the robustness of our results by combining a propensity score matching (PSM) method with DID and using alternative measures and specifications for home sharing's entry.

Our analyses glean important insights. We first investigate how management responses change across hotels after home sharing's entry. Although the average responses remain stable, we find a sharp divergence between higher-priced hotels and lower-priced hotels in their response activities. Specifically, management responses to reviews surge by 2.7% at higher-priced hotels while plummeting by 3.2% at lower-priced hotels. This heterogeneity seems to show the distinct reactions between hotel price segments, electing to fight (by higher-priced hotels) versus retreat (by lower-priced hotels) when facing home-sharing rivals.

What do the asymmetric responses mean to incumbents' performance? We further investigate the change in customer satisfaction and sales across hotels after home sharing's entry. Quite interestingly, we find higher-priced hotels enjoy a rise in both satisfaction (proxied by the average review rating) and sales (proxied by the number of purchase-verified reviews[4]). In contrast, both satisfaction and sales drop for lower-priced hotels. Nevertheless, the divergent performance appears to be largely driven by management responses, regardless of hotel characteristics (e.g., price segment). Although lower-priced hotels overall respond less to reviews after home sharing's entry, those striving to respond more gain higher performance. That is, hotels in any price segment could essentially improve satisfaction and sales if proactively adjusting quality by responding to and learning from reviews. Additionally, management responses seem to not only defend a hotel from external disruption (home sharing) but also differentiate the hotel from its internal peers (which respond less), resulting in increased satisfaction and sales performance.

If management responses are effective, what performance-improving opportunities that hotels have exactly learned by responding to reviews?  We exploit the richness in our data to unveil the content features (topics and sentiments) of review texts. Because each review is

[4] Because not all the customers who have stayed at a hotel would leave a review and the maximal number of reviews associated with each verified purchase is capped at one by the review platform, our definition of sales implies a conservative estimate of the management responses' impact after home sharing's entry.

essentially “a bag of topics,” we first use natural language processing’s LDA (Blei, Ng, and Jordan 2003) to extract theme-specific topics from the massive sentence-level hotel reviews. These topics with the highest coherence score include the check-in/out process, facility and amenity, cleanliness, excursion opportunity, room condition, location, and customer service. We then utilize deep learning’s Word2Vec and CNN algorithms to evaluate the sentiment associated with each topic. We replicate the same set of machine learning procedures on home sharing’s reviews. The results suggest that cleanliness, check-in/out process, room condition, and excursion opportunity are the top four topics that speak to the quality gaps between hotels and home sharing.

We next formally test whether these topics, if responded and learned by hotel managers, would explain the improved hotel performance after home sharing’s entry. We model the impact of entry on the sentiments of seven topics and find hotels respond to reviews more indeed bridge the gap by obtaining higher sentiments on these four topics. Using management responses, hotel managers appear to have learned areas that are not yet up to the game and need to be improved. Such learning is effective and rewarding: hotels achieve higher satisfaction and sales as a way to act to home sharing’s disruption. Mining the granularity and informativeness of unstructured reviews, our machine learning algorithms, along with quasi-experimental evidence, explain why highly responsive hotels tend to gain higher performance. This finding confirms our earlier conjecture: viewing home sharing solely as a threat to incumbents might oversimplify the phenomenon. Rather, incumbents can proactively seek opportunities in rivals’ entry using management responses. The opportunity, based on our findings, is to adjust quality provision in areas such as cleanliness, excursion opportunity, room condition, and check-in/out process to bridge the gap with home sharing on the quality provision.

Our research provides several important implications. This paper first extends the literature on incumbents and home sharing's entry, management responses, and machine learning on review texts. Specifically, our study makes one of the first attempts to investigate if and how incumbent firms react to sharing economy's disruption using management responses to improve quality provision. The findings reveal a unique duo-effect of incumbents' reaction in not only defending external disruptors but also differentiating from internal incumbent peers. Our hybrid method integrating causal inference and machine learning also adds to the emerging literature that lies at the intersection of marketing and technology.

The paper also has direct implications for incumbent firms. We advocate a customer-centric approach (using management responses) to exploit performance-improving opportunities and sustain competitive advantage when facing disruption. Through responding to reviews and automating machine learning, different types of incumbents can extrapolate and scale the efficiency for focused actions specific to their improvement needs while remaining competitive in the market. Specific to the hotel industry practitioners, these actions include adjusting quality provision on cleanliness, excursion opportunity, room condition, and check-in/out process.

The rest of the paper is organized as follows. First, we review relevant literature and discuss our contribution to existing research. We then describe the data and action patterns of incumbents after home sharing's entry. We present the analyses and findings of quasi-experiment and machine learning models, along with additional analyses to rule out alternative explanations. We conclude the paper by discussing its implications.

## Related Literature

### *Incumbents and Home Sharing's Entry*

This paper is first related to the literature on the sharing economy. Research on the rise of sharing economy, primarily home sharing, has been proliferating in recent years, with the majority of the literature focusing on how home sharing disrupts incumbent counterparts that offer similar services. Early work, for example, examines how home sharing cut into hotels' profitability (Farronato and Fradkin 2018; Li and Srinivasan 2019; Zervas, Proserpio, and Byers 2017). More recent research studies how home sharing competes with local housing and rental markets for home supply (Barron, Kung, and Proserpio 2018; Chen, Wei, and Xie 2019; Horn and Merante 2017).

Despite much is known about the sharing economy's impact on incumbents, very little is known about incumbents' responses to sharing economy. Studies such as Wallsten (2015) studying how the incumbent taxi industry reduces customer complaints in response to the growing popularity of Uber are rare. We echo Wallsten (2015) and propose that it is possible sharing economy represents a challenge to the status quo of incumbent services and forces them to adjust quality from what they used to offer. By presenting the first direct empirical evidence on how incumbent businesses adjust quality provision after home sharing's entry, we add to the literature a less researched perspective. Not only do we unveil if and how incumbents respond but also to what extent incumbents can successfully defend for improved performance facing increased competition. Following Zervas, Proserpio, and Byers (2017), we specifically differentiate the responses and performance outcomes across heterogeneous incumbents (hotels), with a focus on the higher-priced and the lower-priced. The results allow us to recommend individualized, responsive tactics for different types of hotels in battling against home sharing.

The paper also contributes to the literature on entry and incumbent firms. From theoretical and empirical analyses, we note there is a vibrant and growing literature on the relationship between entry and price (e.g., McCann and Vroom 2010; Prince and Simon 2015; Seamans 2013). It is intuitive to understand that prices generally fall in the face of increased competition, as prior work has shown. When consumers have more choices, incumbents have an incentive to take steps to become more appealing; perhaps the most obvious step is to reduce the price. We extrapolate from this finding and argue that incumbents tend to compete along additional dimensions (other than price) when facing competition. In particular, we expect that incumbents would want to improve service quality (without costing their profitability bottom-lines) to gain customer satisfaction and sales using management responses. By studying how incumbent adjust their quality provision in response to entry, we suggest a less researched competitive tool (management response for service improvement) to the literature and demonstrate its value to incumbents facing a competitive move such as entry by rivals.

***Management Responses***

As evidenced by the explosion of research on online word-of-mouth, businesses have strong incentives to engage consumers online (Xie, Zhang, and Zhang 2014), learn from the "wisdom of online crowds" (Park and Allen 2013), and determine what they should do to meet customer needs (Kumar, Qiu, and Kumar 2018). Research documents evidence that management responses improve online reputation proxied by the valence of subsequent reviews (Chevalier, Dover, and Mayzlin 2018; Gu and Ye 2014; Wang and Chaudhry 2018), stimulate customer engagement by soliciting customer opinion manifested in the larger volume of reviews (Chen, Gu, Ye, and Zhu, 2019; Proserpio and Zervas 2017), and eventually drive firm performance (Kumar, Qiu, and Kumar 2018).

We add to the literature by advocating another value of responding to reviews: identifying opportunities to fight against disruptors. Without leaving the mechanism behind the impact of management responses in a black box, we provide a clear picture of what exactly the manager has learned from these reviews, and how the manager resorts to the learning to improve performance and compete with the rivals. To this end, we not only measure the sheer number of management responses as a measure of manager efforts, as previous literature does, but also drill down the performance-improving opportunities associated with these responses, assisted by the state-of-the-art machine learning. This represents a contribution of the paper, as our findings speak precisely on areas that the incumbent should focus on when facing increased competition. Additionally, we are also the first to compare and contrast incumbents when embracing the management response as a defensive strategy against disruption. The result sheds light on how proactive and passive responses lead to distinct performance outcomes, adding a less studied perspective on responses heterogeneity to the management response literature.

***Machine Learning on Review Texts***

Research on user-generated content (or reviews) has flourished in the past decades (see, e.g., Chevalier and Mayzlin 2006; Godes and Mayzlin 2004; Liu 2006; and Zhang and Dellarocas 2006). Among them, many have found a positive relationship between reviews and firm performance but only reply on the count (i.e., volume) and the numerical rating (valence) to represent information from reviews. Only until recently, marketing scholars have begun developing methods to mine unstructured textual data to address business questions. In particular, a handful of studies adopt natural language processing to marketing applications (Archak, Ghose, and Ipeirotis 2011; Decker and Trusov 2010; Lee and Bradlow 2011; Tirunillai and Tellis 2014). However, these applications either rely on hand-coded features or engage in intensive feature engineering, which is ad-hoc and error-prone (Liu, Lee, and Srinivasan 2019). The sheer volume of reviews in the "big data" era makes human coding time-consuming.

We join a few pioneer studies in marketing (e.g., Liu, Lee, and Srinivasan 2019; Timoshenko and Hauser, 2019; Zhang and Luo, 2018) to adopt state-of-the-art techniques to extract the content information from the textual review data and achieve scalability. Specifically, we rely on natural language processing using topic modeling (Blei, Ng, and Jordan 2003) and the latest development in deep learning (LeCun, Bengio, and Hinton 2015) that can use raw data to automatically discover feature representations and deal with large-scale, unstructured content information at the highly granular level. Our approach can readily scale up to extract sentence-level topics and sentiments from millions of reviews within hours. To extract the topics and sentiments of review texts at the same time, our approach integrates topic modeling and deep learning using sentences as the basic unit of learning and achieves good performance in the validation process. While machine learning techniques only started to soar to new heights in recent years, we make a meaningful early effort in advocating machine learning for marketing research.

## Data

Different from the domestic (U.S.) sample in most sharing economy literature (e.g., Farronato and Fradkin 2018; Li and Srinivasan 2019; Zervas, Proserpio, and Byers 2017), our data is collected from Beijing, China, a vibrant international destination with growing popularity among travelers as well as hotels, and more recently, home-sharing services. This accommodation market has witnessed an influx of home sharing into the crowded hotel territory between 2015 and 2017, which is our study period (from March 2015 to December 2017, a total of 34 months). Figure W4.1 in Web Appendix W4 presents the distribution of hotels and home-sharing properties in Beijing at the end of our study period.

Our data on home-sharing services is collected from Xiaozhu (xiaozhu.com), the largest peer-to-peer home-sharing platform in China (also known as “Airbnb in China”). As of September 2018, Xiaozhu has more than 8 million home-sharing properties over 400 cities in China.[5] We label the inception of each home-sharing property using the date of its first guest review published on the platform. We identify a home-sharing property’s entry by observing its monthly presence within a 500-meter radius of a hotel.[6]

We collected data on hotels from Ctrip (ctrip.com), the largest online travel agent in China (and the second-largest online travel agent in the world based on valuation).[7] For each hotel in Beijing, we collected its purchase-verified reviews, which are in a hybrid format of numeric ratings (on a scale of 1 – most unsatisfied to 5 – most satisfied) and paragraphs of text. Hotel managers registering a business account with Ctrip can respond underneath the reviews they aim to address. We calculate the average ratings of reviews (a proxy for customer satisfaction), the number of the purchase-verified reviews (a proxy for sales), and the ratio of the number of management responses to the number of reviews (because managers’ responses are proportional to reviews) by hotels and year-month, which is the unit of our analysis. We also store the textual reviews for information extraction in machine learning. In addition to reviews and responses on hotels, we also collected hotel characteristics such as nightly price.

We combine data from different sources into a panel at the hotel by year-month level. Our sample includes 6,678 hotels and 3,584 home-sharing properties in Beijing over a period from March 2015 to December 2017. Table 1 presents the variable definitions and summary

[5] Source: http://www.xiaozhu.com/aboutus.
[6] We also use other radii such as 1,000 meters to define a hotel’s vicinity for robustness checks. The estimation results are consistent with the main specification using the 500-meter radius. We report the analyses in Appendix W3.
[7] Source: http://pages.ctrip.com/public/ctripab/abctrip.htm.

statistics. We take logarithms of several variables (e.g., *Sales*) that exhibit skewed distributions. Therefore, some of our estimations will inherit the semi-elasticity interpretation.

---- Insert Table 1 about here ----

## Analyses and Findings

We present our empirical findings of the incumbent hotels' management responses, the associated performance outcomes, and mechanisms behind the effects in this section. As a quick roadmap, we begin with reporting the specification of the empirical model (i.e., quasi-experiments using DID[8]) and results on the impact of home sharing's entry on hotels' management responses and the performance outcomes (i.e., customer satisfaction and sales), with a focus on the heterogeneous effects across hotels. We then introduce our natural language processing and machine learning procedures and report associated results. Integrating quasi-experiments and machine learning, we also drill down the mechanism behind the improvement in hotel quality provision using management response when facing home sharing's entry. The last section presents additional analyses to rule out alternative explanations of the mechanism. For easy reference, Figure 1 illustrates our research framework, which depicts the flow of analyses.

---- Insert Figure 1 about here ----

### *Impacts of Home Sharing on Hotels' Management Responses and Performance*

The vast variations in the temporal rate of home sharing's entry and geographical expansion across hotel vicinities provide a quasi-experimental opportunity. The sample of our main analysis consists of 4,432 unique hotels that experience home sharing's entry by the end of our sample (constituting the treatment group), and the remaining 2,246 hotels do not experience

[8] We also use a propensity score matching method with DID to check the robustness of our main specification using DID. The result is consistent with the main findings, as reported in Appendix W1.

home sharing's entry throughout the study period (constituting the control group). The main specification of our quasi-experiment is a regression-adjusted difference-in-differences model (Angrist and Pischke 2008). For each hotel *i* in year-month *t*, its management responses are a function of home sharing's entry and other covariates. The equation is as follows:

$$Y_{it} = \beta_0 \cdot Treated_i + \beta_1 \cdot Treated_i \times logSupply_{it} + \mu_i + \nu_t + \delta \cdot Z_{it} + \varepsilon_{it}, \quad (1)$$

where $Y_{it}$ is the dependent variable. Depending on the specific regression, we examine the management responses and performance outcomes (satisfaction and sales) as a result of home sharing's entry. $Treated_i$ identifies whether a hotel has experienced home sharing's entry by the end of the study period. Following the literature (Zervas, Proserpio, and Byers 2017), we use the logarithm of the number of home-sharing properties that have entered the 500-meter radius of hotel *i* in year-month *t*, $logSupply_{it}$, as the treatment.[9] We consider the 500-meter radius[10] a useful a useful geo-boundary for the consideration of market competition between hotels and home sharing based on public and alternative transportation coverage in Beijing. We include hotel fixed effects, $\mu_i$, and year-month fixed effects, $\nu_t$, as in standard DID designs. Note that $Treated_i$ will be absorbed by the hotel fixed effects. Because the measure of home sharing's entry, $logSupply_{it}$, is both hotel and year-month specific, $Treated_i \times logSupply_{it}$ is equivalent to $logSupply_{it}$.[11] Hence, we keep only $logSupply_{it}$ when reporting results, and its estimated

[9] We also check the robustness of the results using a dummy variable, $Entry_{it}$, to indicate whether home sharing has entered the 500-meter vicinity of a hotel. The estimation reveals qualitatively consistent findings with our main specification and is reported in Appendix W2.

[10] We also use alternative vicinity such as a 1,000-meter radius to check the robustness of our results. The estimation reveals results consistent with our main specification and is reported in Appendix W3.

[11] To see this, both $Treated_i \times logSupply_{it}$ and $logSupply_{it}$ have values of zero in the control group. In the treatment group, they are both zeros before home sharing's entry because $logSupply_{it}$ is zero in this period. After home sharing's entry, both are equal to $logSupply_{it}$ because $Treated_i$ is always one for the treatment group and $Treated_i \times logSupply_{it}$ becomes $logSupply_{it}$.

coefficient captures the home sharing's entry effect on the specific dependent variables describing hotels' responses and subsequent performance.

In the set of control variables ($Z_{it}$), we first include the cumulative historical performance $CumRating_{it-1}$ and $logCumReviews_{it-1}$ for each hotel. The purpose is to control for any inherent quality difference of the hotels. Note that we lagged these two variables to avoid including the current values of the dependent variables. We also control for $logHotels_{it}$, which is the number of other hotels in the vicinity of hotel *i* in year-month *t.* Arguably, the agglomeration and competition from peer hotels are associated with a hotel's reaction and performance. Lastly, the residual term, $\varepsilon_{it}$, captures idiosyncratic shocks.

***Management responses to home sharing's entry.*** We first test whether and how hotels' management responses change when facing home sharing's entry. Table 2 reports the regression results based on our main specification. The dependent variable is the ratio of management responses to reviews, $MRRatio_{it}$. First, we find no significant association between home sharing's entry and management responses based on the entire sample of hotels, as shown in Column (1). On average, hotels appear not significantly responsive to home sharing's entry. We further incorporate the heterogeneity of hotels (by price segment) by interacting $logSupply_{it}$ with $HighPrice_i$, a dummy variable indicating a higher-priced hotel based on whether its price is above the median of all hotels.[12] The result exhibits strongly asymmetric responses between different types of hotels. As Column (2) shows, we find an obvious surge in management responses by higher-priced hotels after home sharing's entry, whereas the management responses plummet in lower-priced hotels. Numerically, on average, a 1% increase in home sharing's

[12] The median price is 238 in our sample. Note that the main effect of $HighPrice_i$ itself is absorbed by the hotel fixed effect because it is time-invariant.

properties leads to a 2.7% increase in a higher-priced hotel's management response ratio. However, there is a 3.2% drop in the management response ratio by a lower-priced hotel.

---- Insert Table 2 about here ----

***Hotel performance after home sharing's entry.*** After observing the divergent responses across hotels, a natural question is whether such divergence leads to different performance outcomes. We next examine how management responses affect hotels' customer satisfaction and sales after home sharing's entry. We present the estimation results in Table 3. As Columns (1) and (3) show, on average, home sharing's entry does negatively impact hotels' customer satisfaction and sales, even though the effects are not statistically significant.

---- Insert Table 3 about here ----

We find, however, a clear divergence in performance outcomes between higher-priced hotels and lower-priced hotels, similar to what we find in Table 2 for management responses. Columns (2) and (4) reveal the difference in satisfaction and sales, respectively, between two hotel price segments through the coefficient of the $logSupply_{it} \times HighPrice_i$ interaction. Specifically, as shown in Column (2), we find customer satisfaction of higher-priced hotels increases significantly while that of lower-priced hotels decreases after home sharing's entry. On average, as the home sharing's properties increase by 1% in a lower-priced hotel's vicinity, its customer satisfaction decreases significantly by about 0.019 (out of 5). Oppositely, a 1% increase in home-sharing properties around a higher-priced hotel drives its customer satisfaction by 0.008.

The hotel sales exhibit a similar pattern in Column (4), with a 1% increase of home-sharing properties in a lower-priced hotel's vicinity associated with 0.037% decrease in sales. This is in line with Zervas, Proserpio, and Byers (2017)'s estimates on the revenue cut of hotels by home sharing's entry. Yet, for higher-priced hotels, the sales increase by 0.013% (= 0.05 - 0.037) against home sharing's disruption.

Taken together, these findings imply higher-priced hotels serve customers better and make more sales after home sharing's entry. The divergence in hotels' management response strategies may have led to the difference in their performance, which we test empirically in the next section.

***The driver of hotel performance after home sharing's entry: Management Responses.*** To further show that the hotel performance is driven by management responses rather than other hotel characteristics (e.g., price segment), we test whether a higher level of management response activities after home sharing's entry indeed leads to higher satisfaction and sales. To this end, we first develop a new measure, $HighMR_{it}$, which indicates whether a hotel's management response ratio from the previous month is indeed higher than its value before home sharing's entry.[13] In other words, $HighMR_{it}$ measures whether a hotel has actually increased the efforts to learn from reviews. We then interact $logSupply_{it}$ with $HighMR_{it}$ and add it to Equation (1).[14] If the management response strategy (rather than the price segment of hotels) indeed drives up satisfaction and sales, we should expect a positive and significant coefficient for the interaction term $logSupply_{it} \times HighMR_{it}$. To incorporate the heterogeneity of hotels, we estimate the interaction term in both higher-priced and lower-priced subsamples.

---- Insert Table 4 about here ----

[13] Because the management response ratio in the current month is related to the review number in the month, we use the management response ratio from last month to avoid simultaneity.

[14] Because $HighMR_{it}$ can only be one for the treatment group after home sharing's entry, it implicitly captures the effect of home sharing's entry (equivalent to $Entry_{it} \times HighMR_{it}$). Hence, we only include the moderation term $logSupply_{it} \times HighMR_{it}$ in the specification.

We present the results in Table 4. First, Columns (1) and (2) report the moderating effect of $HighMR_{it}$ on customer satisfaction. Column (1) shows that while home sharing's entry alone does not significantly affect a higher-priced hotel's customer satisfaction, the hotel has an increase in satisfaction if it has a higher management response ratio. This suggests that the positive effect of home sharing's entry on higher-priced hotels' customer satisfaction, which we have observed earlier, seems mainly explained by their higher management response ratio.

Column (2) presents the results in the lower-priced hotel subsample. We find that lower-priced hotels indeed experience a drop in satisfaction facing home sharing's entry. Interestingly, the coefficient of $logSupply_{it} \times HighMR_{it}$ is significantly positive, which seems to suggest that lower-priced hotels that respond more to reviews against the disruption have improved satisfaction. Moreover, we notice this moderation effect in lower-priced hotels is even larger than that in higher-priced hotels, suggesting that lower-priced hotels may benefit even more if they learn and improve using management responses. However, as we have shown earlier, lower-priced hotels, on average, reduce their management response effort (while higher-priced hotels have not), which may have led to the distinct outcomes in customer satisfaction.

Columns (3) and (4) of Table 4 show a similar pattern of results on hotel sales. Consistent with the literature, we find that home sharing's entry has a more significant negative impact on the sales of lower-priced hotels than that of higher-priced hotels. The positive moderation effect of higher management response ratio, though, is larger in lower-priced hotels. That is, a lower-priced hotel would have attenuated the negative impact of home sharing's entry if it responded to and learned from reviews. This divergence further differentiates the sales performance between higher-priced and lower-priced hotels.

Overall, our results suggest that, when facing home sharing's entry, some hotels have actively responded to reviews and identified from reviews the performance-improving opportunities. Their responses appear effective in elevating customer satisfaction and sales against the increased competition from home sharing. Distinct management responses explain the performance gap between higher-priced hotels and lower-priced hotels. Learning from reviews using management responses to adjust quality provision seems helpful in defending incumbents from home sharing's entry. Next, we examine what exactly managerially responsive hotels have learned from reviews using machine learning.

***Machine Learning on Reviews Texts***

In this section, we adopt the latest development in natural language processing and deep learning to understand what opportunities the hotels have learned from reviews. Each review is essentially a "bag of topics." The goal is to extract the topics of each review and the sentiments associated with these topics. The topics represent different service quality dimensions of the hotels (as reviewed by customers), and the sentiments reflect customers' evaluation on these quality dimensions. The challenge is that even though we have both the text and numerical rating for each review, the sentiments for the topics embedded in each review are latent and not readily observable. To address this challenge, we develop an integrated framework to evaluate the sentiments of review topics using natural language processing and deep learning. Specifically, we first use Latent Dirichlet Allocation (Blei, Ng, and Jordan 2003) to extract topics in each review (Puranam, Narayan, and Kadiyali 2017; Tirunillai and Tellis 2014). Then we adopt Convolutional Neural Networks (dos Santos and Gatti 2014; Kim 2014; Zhang, Zhao, and LeCun 2015) to classify the sentiments of each sentence in reviews.

Figure 2 illustrates our integrated process of topic and sentiment learning. Our approach is to learn the topics and sentiments of each sentence in a review and aggregate them back to the review level. The general idea is to train the LDA and CNN models on the review level, and then classify the topic and sentiments of sentences using the trained model. We also partially verify the aggregated results using sub-ratings (e.g. location, facility, cleanliness, and service) scored by customers as a default option when submitting their reviews to Ctrip. The process contains the following steps:

---- Insert Figure 2 about here ----

(1) Text preprocessing. In this step, we clean up the raw review texts and prepare for the topic and sentiment learning. We first remove the stop words that are not informative about hotels and their service quality. They are usually used for connection instead of for meaning (e.g., “the,” “and,” “when”). Second, because the text is in Chinese, which is character-based, we segment the text into words and apply part-of-speech tagging to retain only words that are nouns, adverbs, or adjectives,[15] which are informative about the hotels and their service quality (Tirunillai and Tellis 2014). Third, we tokenize the words into vectors by mapping each word into a unique number. After this process, each review becomes a list of tokens which represent the original words.

(2) Train LDA and CNN models. In this step, we train the LDA and CNN models using review level data. Because each model contains a few technical details, we describe the training and validation of each model in the next two sections. At the end of this step, we obtain an LDA

[15] The Chinese text segmentation package we use is Jieba, available at https://github.com/fxsjy/jieba.

model that can classify the topic of a given document and a CNN model that can classify the sentiments for a piece of text.

(3) Split reviews. We then break each review into individual sentences by the presence of characters or symbols that signal the end of the sentence (e.g., ".", "?", "!", and the new paragraph character).

(4) Extract sentence topics and sentiments. We classify each sentence using the trained LDA and CNN models. For each sentence, the LDA model provides the probabilities of the sentence belonging to each of the topics. For example, if the LDA model has seven topics, the classification will predict that the sentence belongs to topic 1 with probability 20%, topic 2 with probability 15%, etc. We then use the topic that has the highest probability as the topic of the sentence. Meanwhile, the CNN model predicts a positive number between 0 (representing negative) and 1 (representing positive) as the predicted sentiment for the sentence. We then use this number as the sentence's sentiment score. After this step, we have obtained a pair of topic and sentiment for every single sentence.

(5) Aggregate review topics and sentiments. The reviews in our data set have varying lengths. Hence, a review may contain one sentence or multiple sentences. For each review, we calculate the sentiment for each topic by averaging the sentiment scores on the topic from all sentences in the review (some topics may be missing because no sentence in the review belongs to the topic). We then aggregate the review topics and related sentiments to the hotel-month level and incorporate them into our econometric model to investigate what areas the hotels have addressed and learned from reviews using management responses. Before that, we will briefly introduce the LDA and CNN models in the next two sections.

***Latent Dirichlet Allocation for review topics***. We use probabilistic topic models from natural language processing to extract review topics. The LDA model is the most widely used topic model (Blei, Ng, and Jordan 2003). A topic is defined as a latent distribution over a vocabulary of words that customers use to describe their experience with the hotel. LDA then views a document (review) as a collection of words drawn from one or more topics. In our context, a review exhibits different proportions of the topics by using words from these topics that reflect customer experience. For example, a customer may devote 40% of the review to the facility of the hotel, 30% to its location, and 20% to its check-in/out process. Using an unsupervised probabilistic model and Bayesian inference, LDA infers a predefined number of topics as latent variables from the observed distribution of words in each document. We use the LDA implementation in the MALLET package to generate different topics (McCallum 2002).

Because LDA infers a given number of topics, we need to determine how many topics to keep in our model. The number of topics is usually determined by model selection criteria such as coherence scores (Röder, Both, and Hinneburg 2015). We run the model using 3-20 topics and find that the LDA model yields the highest coherence score when the number of topics is equal to 7. Therefore, we set the number of topics to seven in our LDA model. The top 10 words in each of the topics are presented in Table 5.

---- Insert Table 5 about here ----

We label the seven topics according to the top 10 words that belong to the topic. As Table 5 shows, these topics, by meaning, are Check-in/out process (T1), Facility & amenity (T2), Cleanliness (T3), Excursion opportunity (T4), Room condition (T5), Location (T6), and Customer service (T7). Note that our data have four sub-ratings scored by customers as a default option when submitting their reviews: location, facility, cleanliness, and service. The extracted seven topics contain these four dimensions, which gives us more confidence about the topic extracting process. The LDA model seems to extract more granular dimensions of the service aspects in Topics 1, 4, and 7, as well as the facility aspect in Topics 2 and 5. We later use the matched topics to validate the topic sentiments from our integrated machine learning framework.

***Deep learning for topic sentiments.*** To uncover the sentiments of the topics from review sentences, we adopt CNN to classify the sentences (dos Santos and Gatti 2014; Kim 2014; Zhang, Zhao, and LeCun 2015). The word-based CNN model we used to uncover sentence sentiment is a slight variant of the CNN architecture of Kim (2014).

Figure W4.2 in Web Appendix W4 shows the architecture of our deep learning method. As we have explained earlier, we tokenize the words and use the token vectors as the input of our model. Because our reviews are in Chinese, and we cannot use the existing word embeddings such as Google's Word2Vec, we first add a word-embedding layer in our deep learning model. Essentially, the word-embedding layer takes the text corpus and transforms it into word vectors while preserving semantic distances between the words as much as possible. Each word is represented by a 256-dimensional vector. Then, the word-embedding representations go through a convolution layer, a max-pooling layer, and a fully connected layer to extract features from the review sentences. Finally, the extracted features are used to predict the sentiments of the sentences, which is a score between 0 and 1.

To train the model, we dichotomize ratings of reviews by grouping reviews with a higher rating than 4.2 (the sample mean) as positive (donated as 1) and reviews with ratings lower than 4.0 as negative (donated as 0). We discard reviews with ratings in-between to clearly distinguish positive reviews from negative reviews. To facilitate efficient parameter calibration, the training data include a balanced set of 100,000 negative reviews and 100,000 positive reviews. We use

75% of the reviews as the training sample and the rest 25% as the hold-out testing sample. The CNN model yields an 88.7% accuracy on the testing sample.

We then use the trained model to predict the sentiment of each sentence in a review. Note that while the output in the training data is dichotomized, the actual sentiment score generated for each sentence is a continuous variable between 0 and 1, with a higher score representing a more positive sentiment.

We follow the procedure in Figure 2 to classify the topic and sentiment for each sentence in all reviews and then aggregate the sentiments of different topics back into the review level and then the hotel-month level. Because Ctrip also displays sub-ratings of each customer review on location, facility, cleanliness, and service, we test our classification accuracy by comparing the topic sentiments we obtain to these sub-ratings. We find that the accuracies are 87%, 78%, 75%, and 80% for the four topics, respectively. We then aggregate the sentiments on the seven topics back to the hotel-month level.

Table 6 summarizes the topic sentiments on all hotels. Note that not all hotels have all dimensions in all months. So we lose observations on some topics. We can see that hotel reviews receive highly positive sentiments on facility & amenity (T2), location (T6), and customer service (T7). But hotels do not receive a strong evaluation from customers on other topics.

---- Insert Table 6 about here ----

We then compare the sentiments of commonly reviewed topics between hotels and home sharing. The intuition is to identify which areas home sharing outperforms incumbent hotels and where managers face competitive pressure. We replicate the same natural language processing and deep learning process using reviews on home sharing within the vicinity of hotels. Table 7 presents the comparison between home sharing and hotels by sentiments of review topics. For the ease of reference later, we write the seven topics as T1-T7.

---- Insert Table 7 about here ----

As Table 7 shows, among the seven topics that our algorithms have learned from reviews, the check-in/out process (T1), cleanliness (T3), excursion opportunity (T4), and room condition (T5) are the top four quality dimensions where reviews from home sharing exhibit higher sentiments than hotel reviews. These four topics point to areas for quality adjustment for hotel managers. Next, we investigate whether these areas have been effectively addressed by hotel managers for improved quality using management responses

***Identifying Improved Areas Integrating Machine Learning and Quasi-experiments***

We incorporate machine learning-detected topics and their sentiments, including four topics showing the sentiment gaps between hotels and home sharing, into the DID model. The specification for this estimation is similar to Equation (1), where we replace $Y_{it}$ using the topic sentiments.

---- Insert Table 8 about here ---

The results are presented in Table 8. As we can see, after home sharing's entry, hotels responding less to reviews do not see a significant change in review sentiments on almost all topics (except for the customer service in general). However, for hotels that have adopted a proactive response strategy (with a higher management response ratio), the sentiments on check-in/out process (T1), cleanliness (T3), excursion opportunity (T4), and room condition (T5) have increased, especially the latter three. These four topics, as we have shown earlier, are exactly the areas representing gap quality between hotels and home sharing. Note that the sentiments are between 0 and 1. The magnitude of the coefficients is, therefore, small. The improvement in other dimensions is mostly positive, though not statistically significant.

This finding suggests that, after responding to reviews, hotels have learned performance-improving opportunities through, specifically, improving the check-in/out process, working on the cleanliness and hygiene, providing excursion opportunities, and lifting room condition. Essentially, these dimensions are readily improvable, while other dimensions, such as facility and location, have a much higher cost to adjust. By addressing these areas, hotels appear to have achieved higher satisfaction.

***Ruling out Alternative Explanations***

***Potential change in the reviewer base***. Some may worry that after home sharing's entry, customers who tend to complain have switched from hotels to home sharing. The loss of complaining reviewers for the hotels may have driven customer satisfaction and confounded the proposed impact of management responses. We argue that whether home sharing's entry reduces complaining customers is an empirical question. It may be that hotels now have fewer complaining customers after they switch to home sharing (as someone may worry). The flipside might also be true: home sharing's entry may have introduced to hotels more complaining customers who may critique the hotels' service in comparison with a strong competitor (home sharing). We empirically test the change in the reviewer base to address the debate.

The alternative mechanism (a decrease in complaining reviewers), if true, would predict an increase in the average rating of hotels after home sharing's entry (due to a reduction in reviewers who write negative reviews). Furthermore, such an increase should be larger in lower-priced hotels because they were affected by home sharing's disruption more (Zervas, Proserpio, and Byers 2017. We have already tested whether the average rating of hotels, especially the lower-priced, indeed changes after home sharing's entry. The results by hotel price segments have been reported in Table 3. As Column (1) shows, we do not see a significant increase in the average rating after home sharing's entry. After further dividing the hotels into higher-priced and lower-priced hotels in Column (2), we find a significant satisfaction decrease in lower-priced hotels and an increase for higher-priced hotels. The overall takeaway from Table 3 is that lower-priced hotels experience a significant decrease in satisfaction due to home sharing's entry, which is contrary to the prediction of the alternative explanation.

Following Proserpio and Zervas (2017), we further separate the reviews by the level of rating and check whether home sharing's entry indeed affects the number of negative reviews (e.g., reviews with 1-2 star ratings). The results are presented in Table 9. The upper panel shows that the number of reviews for lower-priced hotels decrease across the board after home sharing's entry. Additionally, negative reviews in Columns (1) and (2) have a lighter decline than more positive reviews in Columns (3) and (4). The moderation effect of $HighMR_{it}$ further reveals that a higher management response ratio from lower-priced hotels mainly increases the number of more positive reviews. Overall, the upper panel shows evidence against the prediction that lower-priced hotels would see a larger decrease in negative reviews.

---- Insert Table 9 about here ---

In the lower panel of Table 9, we present the results for higher-priced hotels. Home sharing's entry alone seems to only affect the more positive reviews, as shown in Column (4). Similarly, a higher level of management responses results in more positive reviews, as shown in Columns (3) and (4). The takeaway of Table 9 speaks on the critical role of management responses as an effective mechanism to explain the relationship between home sharing's entry and the increased reviews ratings. By actively responding to reviews, managers have effectively increased the number of positive reviews.

***A placebo test***. To ensure that it is home sharing's entry driving the results, we conduct a placebo test in which we randomly assign a subset of unaffected hotels as if they were affected by home sharing's entry. Specifically, we focus on only unaffected hotels in this test, and in each draw, we randomly treat 50% of those hotels as if they were subject to home sharing's entry and replicate our main regressions on management response ratio to obtain the coefficients of the home sharing's entry. We repeat this process multiple times in our simulation.

Figure W4.3 in the Web Appendix W4 depicts the distributions of the estimated coefficients from simulations with 1,000 draws each. Panel (a) presents the coefficient for the lower-priced subsample, and Panel (b) displays the coefficient for the higher-priced hotels. As we can see from subsamples, most of the mass (of the distributions) center around zero between -0.01 and 0.01 at worst. This suggests that all the estimated coefficients are not significantly different from zero (in a statistical sense), implying that the "counterfactual" policy constructed in the simulations does not affect the management response ratio. The results from this placebo test lend further support to our findings that it is the home sharing's entry driving the divergence in hotels' management response strategies (between affected and unaffected hotels).

***Robustness Checks***

***Alternative control group: propensity score matching***. To alleviate concerns regarding the comparability between the affected hotels and unaffected hotels, we conduct a robustness check by combining a propensity score matching (PSM, hereafter) method with DID in the Web Appendix W1. We first construct a matched sample with similar attributes on all the observable measures using PSM (Abadie and Imbens 2006). Specifically, we run a logistic regression at the hotel level using the mean values of available observables to predict the probability of a hotel experiencing home sharing's entry. These variables include whether the hotel is higher-priced, its historical satisfaction and sales performance, as well as the number of competitor hotels in its vicinity. In the end, we have 1,171 matched hotels from both the treated and control groups. We present the balance check of covariates of the PSM sample in Table W1.1. The covariates are quite balanced from the matching. There is no significant difference in covariates between the two groups.

We then repeat our analysis on the PSM sample and report the results in Table W1.2. As Column (1) shows, we still find a divergent in terms of the management responses between higher-priced and lower-priced hotels. We also find that higher-priced hotels have higher customer satisfaction and sales after home sharing's entry, as shown in Columns (3) and (5). We also re-examine the difference in the effect of management responses between higher-priced hotels and lower-priced hotels and present the results in Table W1.3. Similarly, we find that the increase in customer satisfaction and hotel sales mainly comes from a higher ratio of management responses, and the moderation effects are higher in lower-priced hotels. The magnitude of the estimates from the PSM sample is also similar to our main results in Tables 2-4. Because the PSM process drops quite a big portion of the sample due to strict matching, we keep the more comprehensive estimates using the full sample as our main results.

***Alternative treatment variables and hotel vicinity***. We also verify the robustness of our finding through a series of additional analyses using alternative measures and specifications. First, instead of $logSupply_{it},$ we use a dummy variable $Entry_{it},$ which indicates whether home sharing has entered a hotel's vicinity, as our treatment variable. The results are reported in Tables W2.1 and W2.2 of the Web Appendix W2. All our estimates remain qualitatively consistent as the main analysis using $logSupply_{it}$, although the magnitude of the coefficients is much larger. This is reasonable because the coefficient of $logSupply_{it}$ captures the elasticity of the dependent variables to the increases in home-sharing properties, while the coefficient of $Entry_{it}$ captures the average increase in the management response ratio, customer satisfaction, and hotel sales after home sharing's entry.

Second, one may worry that the 500-meter radius of a hotel that we have used to calculate home sharing's entry in the main analysis may not be robust. To address this concern, we compute home sharing's entry using a 1,000-meter radius and repeat our analyses. The results, as summarized in Tables W3.1 and W3.2 in the Web Appendix W3, are consistent with our main specification using the 500-meter vicinity. We also notice the magnitude of our estimates is slightly lower than our main results because now the radius is larger and we count more home-sharing properties in the analysis. Overall, our results remain quite robust using different measures and specifications.

## Implications and Concluding Remarks

Sharing economy services are disrupting incumbent firms by cutting into their profits. With rapid and accelerating environmental changes, incumbents need to adapt and react to the competition proactively. In this paper, we explore if and how incumbent firms can quickly align marketing strategies to these changes. On substantive contributions, this paper makes one of the first attempts to investigate if and how incumbent firms react to sharing economy's entry using management responses to adjust quality provision, discover the associated performance outcomes, and unveil the mechanism behind service improvement informed by reviews addressed by management responses. It systematically adds to the existing literature on incumbents and home sharing's entry, management responses, and machine learning on review texts. Meanwhile, it also generates new insights on several issues that are important but less explored in the marketing literature, including how incumbents become adaptable in the face of disruption, the duo-effect of incumbents' reaction in not only defending external disruptors but also differentiating from internal incumbent peers, and how machine learning can be deployed at the business level to assist addressing these questions.

On practical contributions, this paper taps into the external environment of incumbent firms and advocates incumbent firms to quickly align their marketing strategies to new environments as the pace of environmental change accelerates. Despite the competitive landscape constantly driven by technology advancement, we advocate the incumbents should learn to disrupt by taking a customer-centric approach of responding and listening. One of the marketing strategies advocated in this study is for managers to respond to and learn from the ubiquitously available customer reviews for areas to improve – especially areas in which a large quality gap between your business and the disruptors exists. We show how incumbents can utilize management responses to exploit performance-improving opportunities in customer feedback and sustain competitive advantage facing increased competition from home sharing. Through responding to reviews and automating machine learning, different types of incumbents can extrapolate and scale the efficiency for focused, targeted, and actionable strategies specific to their improvement needs while remaining competitive in the market.

Our hybrid method integrating causal inference and machine learning is theoretically and practically relevant. We first see this study as one of the recent attempts dealing with reviews' unstructured format, rapid speed of generation, and the large volume of information (i.e., the "big data" problems) for marketing research. Despite the impact of reviews on business performance has been well established in the marketing literature (Chevalier and Mayzlin 2006), much less work has focused on insights from review content, beyond review ratings, that speak to the specific performance-improving opportunities. Through systematically transforming reviews to connect directly to economic outcomes, this paper, along with other data-driven, analytics-focused studies, adds to the emerging literature that lies at the intersection of marketing and technology and is making a step towards the MarTech paradigm for the technology-centric future of marketing. Second, the advantage of our machine learning algorithms is that they let us sift these theme-specific content features (topics and sentiments) across a wide range of reviews without hand-coding features with human intervention or domain knowledge. If automated and deployed by hotels, this approach can not only help them extract cues in reviews for performance-improving opportunities but also be used to monitor the customer sentiments for competitors. Leveraging the power of deep learning, the managers can proactively identify areas to bridge the quality gap toward elevated customer satisfaction and sales performance, thus battling back against the disruption from home sharing.

**References**

Abadie, Alberto and Guido W. Imben (2006), "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74(1), 235–267.

Ankur (2018), "Service Quality & Customer Satisfaction In The Hotel Industry," *Trilyo*, (September 2018), (accessed June 21, 2019), [available at https://www.trilyo.com/blog/service-quality-customer-satisfaction-in-the-hotel-industry/]

Angrist, Joshua D. and Jörn-Steffen Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Archak, Nikolay, Anindya Ghose, and Panagiotis G. Ipeirotis (2011), "Deriving the pricing power of product features by mining consumer reviews," *Management Science*, 57(8), 1485–1509.

Barron, Kyle, Edward Kung, and Davide Proserpio (2018), "The Sharing Economy and Housing Affordability: Evidence from Airbnb," *SSRN*, (March 29, 2018), (accessed October 1, 2018), [available at https://ssrn.com/abstract=3006832].

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3(Jan), 993–1022.

Chen, Wei, Zaiyan Wei, and Karen, Xie (2019), "The Battle for Homes: How Does Home Sharing Disrupt Local Residential Markets?" *SSRN*, (June 2019), (accessed June 1, 2019), [available at https://ssrn.com/abstract=3257521].

Chen, Wei, Bin Gu, Qiang Ye, and Kevin Zhu (2018), "Measuring and Managing the Externality of Managerial Responses to Online Customer Reviews," *Information Systems Research*, 30(1), 81-96.

Chevalier, Judith A., Dover Yaniv, and Dina Mayzlin (2018), “Channels of Impact: User Reviews When Quality Is Dynamic and Managers Respond,” *Marketing Science,* 37(5), 685-853.

Chevalier, Judith A. and Dina Mayzlin (2006), “The effect of word of mouth on sales: Online book reviews,” *Journal of Marketing Research,* 43(3), 345–354.

Decker, Reinhold and Michael Trusov (2010), “Estimating aggregate consumer preferences from online product reviews,” *International Journal of Research in Marketing*, 27(4), 293–307.

dos Santos, Cicero and Maíra Gatti (2014), “Deep convolutional neural networks for sentiment analysis of short texts,” *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pp. 69–78.

Farronato, Chiara and Andrey Fradkin (2018), “The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb,” *NBER Working Paper No. 24361*.

Godes, David and Dina Mayzlin (2004), “Using online conversations to study word-of-mouth communication,” *Marketing Science,* 23(4), 545–560.

Gu, Bu and Qiang Ye (2014), “First step in social media: measuring the influence of online management responses on customer satisfaction,” *Production and Operations Management*, 23 (4),570–582.

Horn, Keren and Mark Merante (2017), “Is Home Sharing Driving up Rents? Evidence from Airbnb in Boston,” *Journal of Housing Economics,* 38(December), 14-24.

Kim, Yoon (2014), “Convolutional Neural Networks for Sentence Classification,” *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, EMNLP

2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL. pp. 1746–1751.

Naveen, Kumar, Liangfei Qiu, and Subodha Kumar (2018), "Exit, Voice, and Response on Digital Platforms: An Empirical Investigation of Online Management Response Strategies," *Information Systems Research*, 29(4), 849-870.

Li, Hui and Srinivasan K (2019), "Competitive Dynamics in the Sharing Economy: An Analysis in the Context of Airbnb and Hotels," *Marketing Science*, 38(3), 365-391.

Liu, Yong (2006), " Word-of-mouth for movies: its dynamics and impact on box office receipts," *Journal of Marketing*, 70(3), 74–89.

Liu, Xiao, Dokyun Lee, and Kannan Srinivasan (2019), "Large Scale Cross Category Analysis of Consumer Review Content on Sales Conversion Leveraging Deep Learning," *Journal of Marketing Research*, 56(6), 918-943.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015), "Deep learning," *Nature*, 521(7553), 436–444.

Lee, Thomas Y. and Eric T. Bradlow (2011), "Automated marketing research using online customer reviews," *Journal of Marketing Research*, 48(5), 881–894.

McCallum, Andrew K. (2002), "MALLET: A machine learning for language toolkit," *Umass Amherst Common Public License*, (2002), (accessed June 1, 2019), [available at http://mallet.cs.umass.edu/].

McCann, Brian T. and Govert Vroom (2010), "Pricing response to entry and agglomeration effects," *Strategic Management Journal*, 31(3), 284-305.

Park, Sung-Yung and Jonathan P. Allen (2013), "Responding to online reviews: problem solving and engagement in hotels," Cornell Hospitality Quarterly, 54 (1), 64–73.

Prince Jeffrey T. and Simon, Daniel H. (2015), "Do incumbents improve service quality in response to entry? Evidence from airlines' on-time performance," *Management Science*, 61(2), 372-390.

Proserpio, Davide and Georgios Zervas (2017), "Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews," *Marketing Science,* 36(5), 645-665.

Puranam, Dinesh, Vishal Narayan, and Vrinda Kadiyali (2017)," The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science,* 36(5), 645-812.

Röder, Michael, Andreas Both, and Alexander Hinneburg (2015), "Exploring the Space of Topic Coherence Measures," *WSDM '15 Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pp. 399-408.

Seamans, Robert C. (2013), "Threat of entry, asymmetric information, and pricing," *Strategic Management Journal,* 34(4), 426-444.

Simon, Daniel (2005), "Incumbent pricing responses to entry," *Strategic Management Journal,* 26(13), 1229-1248.

Artem Timoshenko and John R. Hauser (2019), "Identifying Customer Needs from User-Generated Content," *Marketing Science*, 38(1), 1-20.

Tirunillai, Seshadri and Gerard J. Tellis (2014), "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51(4), 463–479.

Wallsten Scott (2015), "The Competitive Effects of the Sharing Economy: How is Uber Changing Taxis?" *Technology Policy Institute Working paper*, (June 2015), (accessed

June 1, 2018), [available at www.ftc.gov/system/files/documents/public_comments/2015/06/01912-96334.pdf].

Wang, Yang and Alexander Chaudhry (2018), "When and how Managers' Responses to Online Reviews Affect Subsequent Reviews," *Journal of Marketing Research*, 55(2), 163-177.

Xie, Karen, Ziqiong Zhang, and Zili, Zhang (2014), "The business value of online consumer reviews and management response to hotel performance," *International Journal of Hospitality Management,* 43(1), 1-12.

Zervas, Georgios, Davide Proserpio, and John W. Byers (2017), "The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry," *Journal of Marketing Research* 54(5), 687–705.

Zervas, Georgios, Davide Proserpio, and John W. Byers (2015), "A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average," *SSRN*, (January 28, 2015), (accessed June 1, 2018), [available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2554500].

Zhang, Xiaoquan and Chris Dellarocas (2006), "The Lord Of The Ratings: Is A Movie's Fate is Influenced by Reviews?' *ICIS 2006 Proceedings*, pp.117.

Zhang, Mengxia and Lan Luo (2018), "Can User-Posted Photos Serve as a Leading Indicator of Restaurant Survival? Evidence from Yelp," SSRN, (March 1, 2018), (accessed June 2019), [available at https://ssrn.com/abstract=3108288].

Zhang, Xiang, Junbo Zhao, and Yann LeCun (2015), "Character-level Convolutional Networks for Text Classification," *Proceedings of the 28th International Conference on Neural Information Processing Systems*, NIPS'15. Cambridge, MA: MIT Press, pp. 649–657.

Table 1. Variable Definitions and Summary Statistics

| | Definition | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|
| ***Hotel performance*** | | | | | | |
| *Satisfaction* | Average rating of hotel reviews | 4.214 | 4.380 | 0.713 | 1 | 5 |
| *Sales* | Logarithm of the number of purchase-verified hotel reviews | 2.073 | 2.079 | 1.400 | 0 | 6.688 |
| ***Home sharing's entry*** | | | | | | |
| *logSupply* | The logarithm of the number of home-sharing properties that have entered a hotel's vicinity | 0.990 | 0.693 | 1.105 | 0 | 4.754 |
| *Entry* | A dummy variable indicating whether home sharing has entered a hotel's vicinity | 0.641 | 1 | 0.480 | 0 | 1 |
| ***Management responses of hotels*** | | | | | | |
| *MRRatio* | The ratio of the number of management responses to the number of hotel reviews | 0.382 | 0 | 0.464 | 0 | 1 |
| *HighMR* | A dummy variable indicating whether a hotel's response ratio in last month is higher than its average response ratio before home sharing's entry | 0.221 | 0 | 0.415 | 0 | 1 |
| ***Hotel characteristics*** | | | | | | |
| *CumRating* | Cumulative average rating of hotel reviews until last month | 4.190 | 4.275 | 0.482 | 1 | 5 |
| *logCumReviews* | The logarithm of the number of cumulative reviews until last month | 4.493 | 4.700 | 1.823 | 0 | 9.265 |
| *logHotels* | The logarithm of the number of other hotels in a hotel's vicinity | 2.239 | 2.398 | 1.015 | 0 | 4.574 |
| *HighPrice* | A dummy variable indicating a higher-priced hotel based on whether its price is higher than the median of all hotels in our sample | 0.495 | 0 | 0.500 | 0 | 1 |

Table 2. Impact of Home Sharing's Entry on Management Responses across Hotels

| D.V.: *MRRatio* | (1) | (2) |
|---|---|---|
| *logSupply* | 0.002 | -0.032*** |
| | (0.002) | (0.003) |
| *logSupply* × *HighPrice* | | 0.059*** |
| | | (0.003) |
| *logHotels* | -0.037*** | -0.040*** |
| | (0.007) | (0.007) |
| *CumRating* | -0.004*** | -0.002** |
| | (0.001) | (0.001) |
| *logCumReviews* | 0.007*** | 0.005*** |
| | (0.001) | (0.001) |
| | | |
| Hotel FE | YES | YES |
| Month FE | YES | YES |
| Observations | 118,111 | 118,111 |
| R-squared | 0.706 | 0.707 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

Table 3. Impact of Home Sharing's Entry on Customer Satisfaction and Sales across Hotels

| D.V.s: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Satisfaction* | | *Sales* | |
| *logSupply* | -0.003 | -0.019*** | -0.008 | -0.037*** |
| | (0.004) | (0.006) | (0.005) | (0.006) |
| *logSupply* × *HighPrice* | | 0.027*** | | 0.050*** |
| | | (0.006) | | (0.007) |
| *logHotels* | 0.014 | 0.012 | -0.051*** | -0.054*** |
| | (0.013) | (0.013) | (0.015) | (0.015) |
| *CumRating* | -0.020*** | -0.020*** | 0.083*** | 0.084*** |
| | (0.003) | (0.003) | (0.002) | (0.003) |
| *logCumReviews* | 0.011*** | 0.009*** | 0.041*** | 0.038*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| | | | | |
| Hotel FE | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES |
| Observations | 118,111 | 118,111 | 118,111 | 118,111 |
| R-squared | 0.350 | 0.350 | 0.819 | 0.819 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

Table 4. Impact of Home Sharing's Entry and Management Responses by Hotels on Customer Satisfaction and Sales

| D.V.s: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Satisfaction* | | *Sales* | |
| | Higher-priced | Lower-priced | Higher-priced | Lower-priced |
| *logSupply* | -0.006 | -0.016** | -0.003 | -0.050*** |
| | (0.005) | (0.008) | (0.007) | (0.007) |
| *logSupply* × *HighMR* | 0.011*** | 0.028*** | 0.032*** | 0.057*** |
| | (0.003) | (0.006) | (0.005) | (0.007) |
| *logHotels* | -0.014 | 0.043** | -0.006 | -0.089*** |
| | (0.014) | (0.021) | (0.020) | (0.021) |
| *CumRating* | -0.018*** | -0.020*** | 0.113*** | 0.067*** |
| | (0.003) | (0.004) | (0.004) | (0.003) |
| *logCumReviews* | 0.006** | 0.009** | 0.054*** | 0.020*** |
| | (0.003) | (0.004) | (0.004) | (0.004) |
| Hotel FE | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES |
| Observations | 58,448 | 59,663 | 58,448 | 59,663 |
| R-squared | 0.362 | 0.313 | 0.835 | 0.738 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

Table 5. Topics and Associated Top 10 Words Mined from Reviews

| Review Topic | Top 10 Words of the Topic |
|---|---|
| Topic 1 (T1): Check-in/out process | front desk, check-in, bellman, attitude, checkout, cannot, guest, booking, problem, telephone |
| Topic 2 (T2): Facility & amenity | hotel, breakfast, feeling, eat, like, children, experience, décor, restaurant, hallway |
| Topic 3 (T3): Cleanliness | room, AC, toilet, window, broken, hot water, dirty, bedding, bed, smell |
| Topic 4 (T4): Excursion opportunity | stay, go, hotel, people, find, look, place, walk, suggestion, garbage |
| Topic 5 (T5): Room condition | room, bad, facility, small, price, big, old, noise, expensive, cheap |
| Topic 6 (T6): Location | good, convenient, close, location, transport, subway, go out, airport, dining, distance |
| Topic 7 (T7): Customer service | good, environment, clean, not bad, price-quality ratio, attitude, satisfied, recommend, comfortable, warm |

Table 6. Summary Statistics of Topic Sentiments of Hotel Reviews

| Review Topic | Obs. | Mean Sentiment | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Topic 1 (T1): Check-in/out process | 73,159 | .461 | 0.290 | 0 | 1 |
| Topic 2 (T2): Facility & amenity | 64,066 | .760 | 0.304 | 0 | 1 |
| Topic 3 (T3): Cleanliness | 60,113 | .351 | 0.337 | 0 | 1 |
| Topic 4 (T4): Excursion opportunity | 64,879 | .596 | 0.346 | 0 | 1 |
| Topic 5 (T5): Room condition | 74,268 | .449 | 0.299 | 0 | 1 |
| Topic 6 (T6): Location | 84,406 | .796 | 0.208 | 0 | 1 |
| Topic 7 (T7): Customer service | 85,647 | .902 | 0.169 | 0 | 1 |

Table 7. Difference between Home Sharing and Hotels by Topic Sentiments

| Review Topic | (1) Hotels (Mean) | (2) Home Sharing (Mean) | (3) Sentiments Difference | (4) Rank by Sentiment Difference |
|---|---|---|---|---|
| Topic 1 (T1): Check-in/out process* | .408 | .71 | .302 | 2 |
| Topic 2 (T2): Facility & amenity | .692 | .922 | .23 | 5 |
| Topic 3 (T3): Cleanliness* | .331 | .651 | .32 | 1 |
| Topic 4 (T4): Excursion opportunity* | .56 | .807 | .247 | 4 |
| Topic 5 (T5): Room condition* | .428 | .716 | .288 | 3 |
| Topic 6 (T6): Location | .771 | .894 | .123 | 6 |
| Topic 7 (T7): Customer service | .876 | .953 | .077 | 7 |

*Note.* T-tests of the mean comparison on all topics are significant at 0.001 level. The top four topics that exhibit the largest quality gap between home sharing and hotels are labeled with *.

Table 8. Impact of Home Sharing's entry and Management Responses by Hotels on Review Topic Sentiments

| D.V.s: Average Sentiments of Topics 1-7 | (1) T1 Check-in/out process | (2) T2 Facility & amenity | (3) T3 Cleanliness | (4) T4 Excursion opportunity | (5) T5 Room conditions | (6) T6 Location | (7) T7 Customer service |
|---|---|---|---|---|---|---|---|
| *logSupply* | 0.000 | 0.000 | -0.000 | -0.002 | -0.001 | 0.000 | 0.004** |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.003) | (0.002) | (0.002) |
| *logSupply* × *HighMR* | 0.004* | 0.001 | 0.007*** | 0.008*** | 0.006*** | 0.002 | -0.000 |
| | (0.002) | (0.002) | (0.003) | (0.002) | (0.002) | (0.001) | (0.001) |
| *logHotels* | 0.019** | 0.017** | -0.006 | 0.002 | -0.001 | 0.008 | -0.004 |
| | (0.008) | (0.008) | (0.011) | (0.010) | (0.008) | (0.005) | (0.004) |
| *CumRating* | -0.004** | -0.000 | -0.002 | 0.000 | 0.002 | 0.001 | -0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| *logCumReviews* | 0.002 | -0.004** | -0.003 | -0.000 | -0.005*** | -0.004*** | -0.002** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| Hotel FE | YES | YES | YES | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES | YES | YES | YES |
| Observations | 73,159 | 64,066 | 60,113 | 64,879 | 74,268 | 84,406 | 85,647 |
| R-squared | 0.193 | 0.228 | 0.200 | 0.227 | 0.220 | 0.174 | 0.174 |

*Note.* Clustered robust standard errors in parentheses.*** $p$<0.01, ** $p$<0.05, * $p$<0.1

Table 9. Impact of Home Sharing's Entry on Hotel Reviews with Different Levels of Ratings

| D.V.s: | (1)<br>log(Reviews w/ Ratings 1 to 2) | (2)<br>log(Reviews w/ Ratings 2 to 3) | (3)<br>log(Reviews w/ Ratings 3 to 4) | (4)<br>log(Reviews w/ Ratings 4 to 5) |
|---|---|---|---|---|
| Panel 1. Lower-priced Hotels | | | | |
| *logSupply* | -0.010** | -0.026*** | -0.041*** | -0.046*** |
| | (0.005) | (0.005) | (0.006) | (0.007) |
| *logSupply* × *HighMR* | -0.002 | -0.000 | 0.015** | 0.049*** |
| | (0.004) | (0.005) | (0.006) | (0.006) |
| Panel 2. Higher-priced Hotels | | | | |
| *logSupply* | 0.004 | -0.003 | 0.004 | -0.019*** |
| | (0.005) | (0.005) | (0.006) | (0.006) |
| *logSupply* × *HighMR* | -0.001 | 0.000 | 0.008** | 0.028*** |
| | (0.003) | (0.003) | (0.004) | (0.004) |

*Notes*. Clustered robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. We only present the coefficients of major treatment variables due to space constraints.

Figure 1. Research Framework

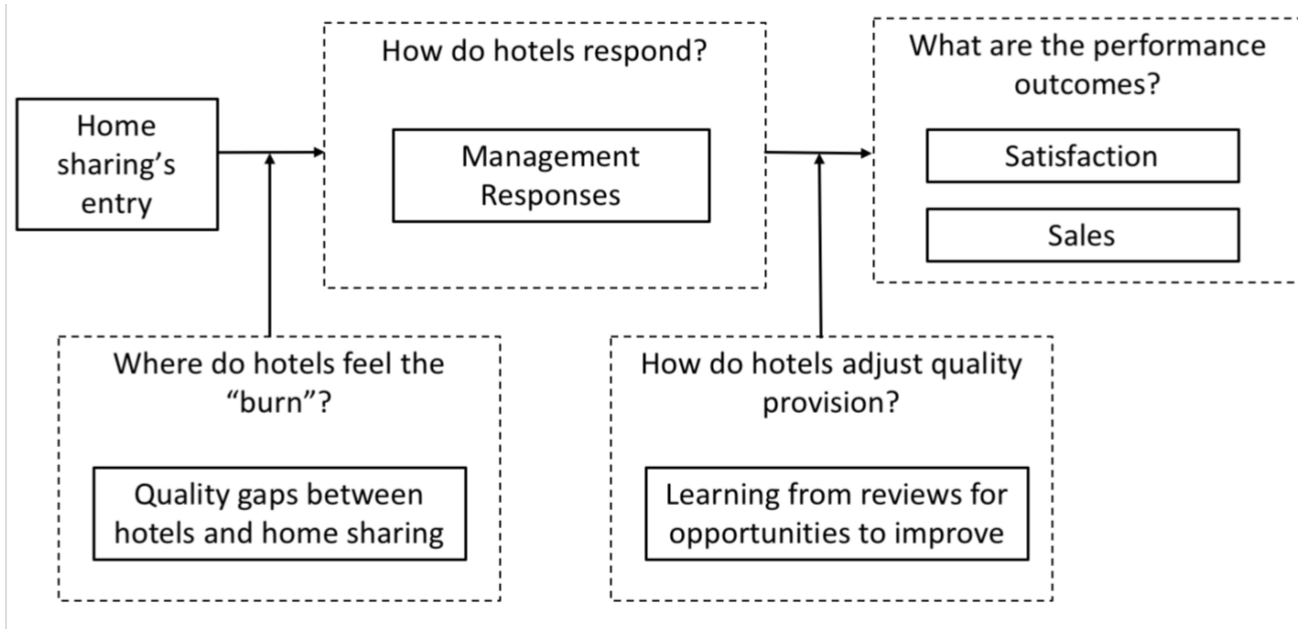
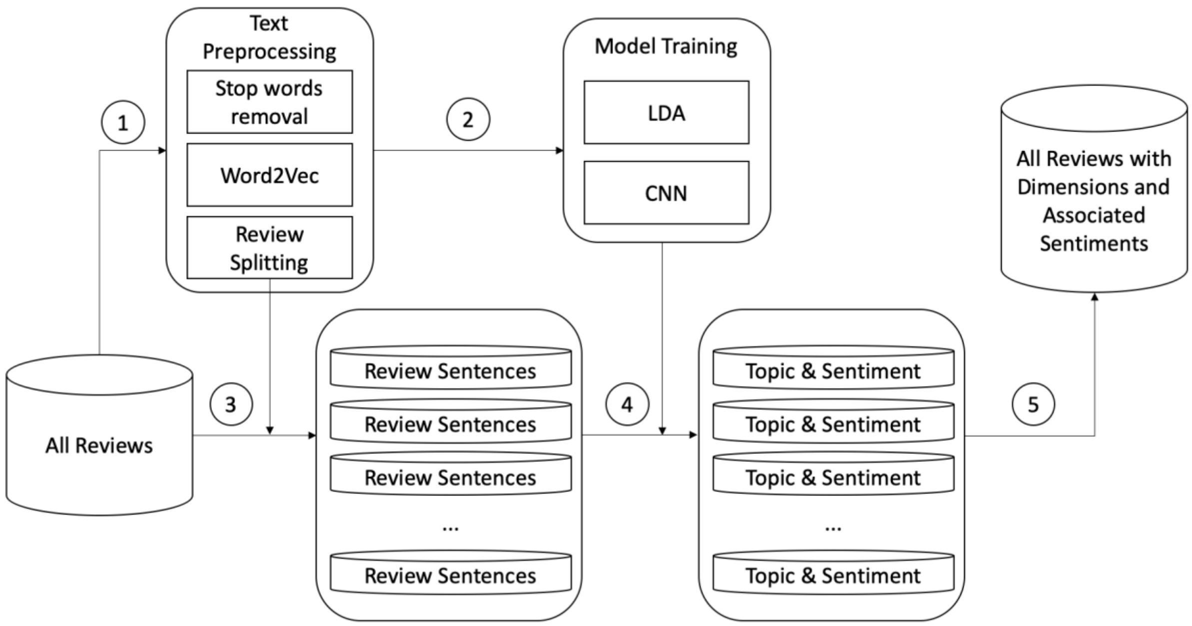


Figure 2. Integrated Machine Learning Process for Reviews Topics and Sentiments



# Web Appendixes

## Appendix W1. Regression Results Using the Propensity Score Matched Sample

Table W1.1. Balance Check of Covariates in the PSM Sample

| Variable | Control | | Treatment | | Difference | | |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Diff | SE | p-value |
| *MRRatio* | 0.387 | 0.435 | 0.386 | 0.435 | -0.001 | 0.023 | 0.969 |
| *Sales* | 1.655 | 1.199 | 1.623 | 1.205 | -0.032 | 0.065 | 0.625 |
| *Satisfaction* | 4.114 | 0.630 | 4.111 | 0.482 | -0.003 | 0.030 | 0.925 |
| *HighPrice* | 0.366 | 0.482 | 0.347 | 0.476 | -0.019 | 0.026 | 0.466 |
| *logHotels* | 1.790 | 0.852 | 1.798 | 0.794 | 0.007 | 0.044 | 0.870 |
| *CumRating* | 2.750 | 1.347 | 2.786 | 1.151 | 0.036 | 0.067 | 0.589 |
| *logCumReviews* | 2.887 | 2.118 | 2.728 | 2.033 | -0.159 | 0.112 | 0.154 |

Table W1.2. Impact of Home Sharing's Entry on Management Responses, Customer Satisfaction, and Sales by Hotels in the PSM Sample

| D.V.s: | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *MRRatio* | *Satisfaction* | | *Sales* | |
| *logSupply* | -0.028*** | 0.007 | -0.004 | 0.008 | -0.007 |
| | (0.004) | (0.006) | (0.009) | (0.007) | (0.009) |
| *logSupply* × *HighPrice* | 0.069*** | | 0.019** | | 0.026** |
| | (0.004) | | (0.009) | | (0.011) |
| *logHotels* | -0.026*** | -0.001 | -0.001 | -0.043** | -0.043** |
| | (0.008) | (0.015) | (0.015) | (0.019) | (0.019) |
| *CumRating* | -0.006*** | -0.015*** | -0.014*** | 0.082*** | 0.082*** |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| *logCumReviews* | 0.008*** | 0.006 | 0.005 | 0.037*** | 0.036*** |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| | | | | | |
| Hotel FE | YES | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES | YES |
| Observations | 58,475 | 58,475 | 58,475 | 58,475 | 58,475 |
| R-squared | 0.690 | 0.307 | 0.308 | 0.784 | 0.784 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

Table W1.3. Impact of Home Sharing's Entry and Management Responses by Hotels on Customer Satisfaction and Sales in the PSM Sample

| D.V.s: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Satisfaction* | | *Sales* | |
| | Higher-priced | Lower-priced | Higher-priced | Lower-priced |
| *logSupply* | 0.003 | 0.001 | 0.015 | -0.022** |
| | (0.007) | (0.011) | (0.010) | (0.010) |
| *logSupply* × *HighMR* | -0.001 | 0.042*** | 0.020** | 0.086*** |
| | (0.006) | (0.011) | (0.009) | (0.015) |
| *logHotels* | -0.015 | 0.014 | 0.021 | -0.098*** |
| | (0.016) | (0.025) | (0.026) | (0.027) |
| *CumRating* | -0.011** | -0.014** | 0.131*** | 0.068*** |
| | (0.005) | (0.005) | (0.009) | (0.005) |
| *logCumReviews* | 0.004 | 0.003 | 0.057*** | 0.016*** |
| | (0.005) | (0.006) | (0.007) | (0.006) |
| Hotel FE | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES |
| Observations | 26,501 | 31,974 | 26,501 | 31,974 |
| R-squared | 0.360 | 0.268 | 0.798 | 0.693 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

## Appendix W2. Regression Results Using the Entry Dummy

Table W2.1. Impact of Home Sharing's Entry on Management Responses, Customer Satisfaction, and Sales by Hotels Using the Entry Dummy

| D.V.s: | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *MRRatio* | *Satisfaction* | | *Sales* | |
| *Entry* | -0.067*** | 0.002 | -0.020* | -0.014 | -0.041*** |
| | (0.005) | (0.008) | (0.012) | (0.009) | (0.012) |
| *Entry* × *HighPrice* | 0.111*** | | 0.045*** | | 0.055*** |
| | (0.007) | | (0.013) | | (0.016) |
| *logHotels* | -0.035*** | 0.012 | 0.013 | -0.054*** | -0.054*** |
| | (0.006) | (0.013) | (0.013) | (0.015) | (0.015) |
| *CumRating* | -0.003*** | -0.020*** | -0.020*** | 0.083*** | 0.084*** |
| | (0.001) | (0.003) | (0.003) | (0.002) | (0.002) |
| *logCumReviews* | 0.006*** | 0.011*** | 0.010*** | 0.041*** | 0.040*** |
| | (0.001) | (0.003) | (0.003) | (0.003) | (0.003) |
| Hotel FE | YES | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES | YES |
| Observations | 118,111 | 118,111 | 118,111 | 118,111 | 118,111 |
| R-squared | 0.707 | 0.350 | 0.350 | 0.819 | 0.819 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

Table W2.2. Impact of Home Sharing's Entry and Management Responses by Hotels on Customer Satisfaction and Sales Using the Entry Dummy

| D.V.s: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Satisfaction* | | *Sales* | |
| | Higher-priced | Lower-priced | Higher-priced | Lower-priced |
| *Entry* | -0.003 | -0.011 | -0.042*** | -0.044*** |
| | (0.008) | (0.014) | (0.013) | (0.013) |
| *Entry* × *HighMR* | 0.029*** | 0.049*** | 0.106*** | 0.148*** |
| | (0.006) | (0.011) | (0.010) | (0.013) |
| *logHotels* | -0.015 | 0.039* | -0.003 | -0.099*** |
| | (0.014) | (0.021) | (0.020) | (0.021) |
| *CumRating* | -0.019*** | -0.021*** | 0.109*** | 0.065*** |
| | (0.003) | (0.004) | (0.004) | (0.003) |
| *logCumReviews* | 0.007** | 0.010** | 0.055*** | 0.020*** |
| | (0.003) | (0.004) | (0.004) | (0.004) |
| Hotel FE | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES |
| Observations | 58,448 | 59,663 | 58,448 | 59,663 |
| R-squared | 0.362 | 0.313 | 0.835 | 0.738 |

*Note*. Clustered robust standard errors in parentheses. ***$p$<0.01, **$p$<0.05, *$p$<0.1

**Appendix W3. Regression Results Using an Alternative Vicinity of Hotels**

Table W3.1. Impact of Home Sharing's Entry on Management Responses, Customer Satisfaction, and Sales by Hotels with its 1000-meter Vicinity

| D.V.s: | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *MRRatio* | *Satisfaction* | | *Sales* | |
| *logSupply* | -0.019*** | 0.005 | -0.008 | 0.003 | -0.019*** |
| | (0.002) | (0.004) | (0.006) | (0.005) | (0.006) |
| *logSupply* × *HighPrice* | 0.049*** | | 0.021*** | | 0.038*** |
| | (0.002) | | (0.004) | | (0.005) |
| *logHotels* | -0.041*** | 0.012 | 0.010 | -0.055*** | -0.057*** |
| | (0.006) | (0.013) | (0.013) | (0.015) | (0.015) |
| *CumRating* | -0.002* | -0.020*** | -0.019*** | 0.083*** | 0.085*** |
| | (0.001) | (0.003) | (0.003) | (0.002) | (0.003) |
| *logCumReviews* | 0.003*** | 0.010*** | 0.009*** | 0.040*** | 0.037*** |
| | (0.001) | (0.003) | (0.003) | (0.003) | (0.003) |
| Hotel FE | YES | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES | YES |
| Observations | 118,111 | 118,111 | 118,111 | 118,111 | 118,111 |
| R-squared | 0.708 | 0.350 | 0.350 | 0.819 | 0.819 |

*Note*. Clustered robust standard errors in parentheses. ****p*<0.01, ***p*<0.05, **p*<0.1

Table W3.2. Impact of Home Sharing's Entry and Management Responses by Hotels on Customer Satisfaction and Sales with its 1000-meter Vicinity

| D.V.s: | (1) | (2) | (1) | (2) |
|---|---|---|---|---|
| | *Satisfaction* | | *Sales* | |
| | Higher-priced | Lower-priced | Higher-priced | Lower-priced |
| *logSupply* | 0.002 | -0.004 | 0.009 | -0.033*** |
| | (0.005) | (0.007) | (0.008) | (0.007) |
| *logSupply* × *HighMR* | 0.007*** | 0.019*** | 0.024*** | 0.042*** |
| | (0.002) | (0.004) | (0.003) | (0.004) |
| *logHotels* | 0.011 | 0.024 | 0.008 | -0.187*** |
| | (0.016) | (0.026) | (0.025) | (0.026) |
| *CumRating* | -0.018*** | -0.020*** | 0.113*** | 0.067*** |
| | (0.003) | (0.004) | (0.004) | (0.003) |
| *logCumReviews* | 0.006* | 0.009** | 0.053*** | 0.019*** |
| | (0.003) | (0.004) | (0.004) | (0.004) |
| | | | | |
| Hotel FE | YES | YES | YES | YES |
| Month FE | YES | YES | YES | YES |
| Observations | 58,448 | 59,663 | 58,448 | 59,663 |
| R-squared | 0.362 | 0.313 | 0.835 | 0.738 |

*Note*. Clustered robust standard errors in parentheses. ****p*<0.01, ***p*<0.05, **p*<0.1

**Appendix W4. Figures for Property Distribution, Deep Learning Model, and Placebo Test**

Figure W4.1. Distribution of Hotels and Home-sharing Properties in Beijing

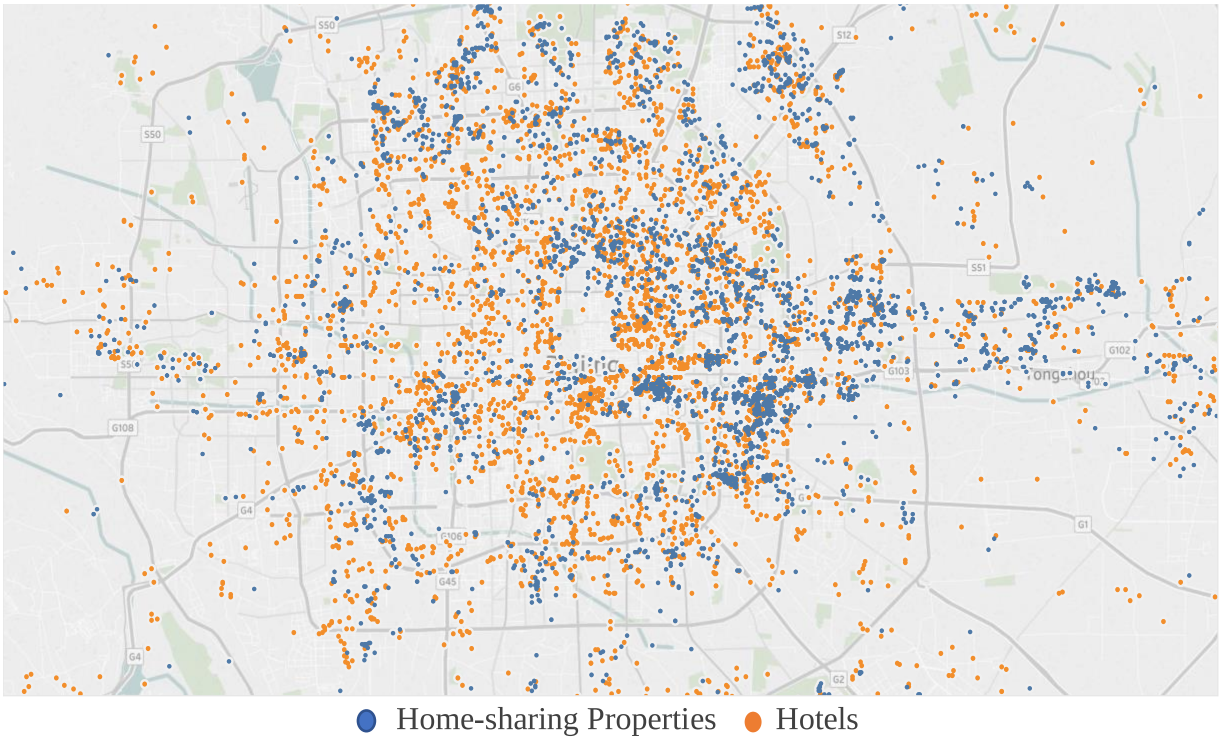


Figure W4.2. Conceptual Structure of Deep Learning Model (Adapted from Kim 2014)

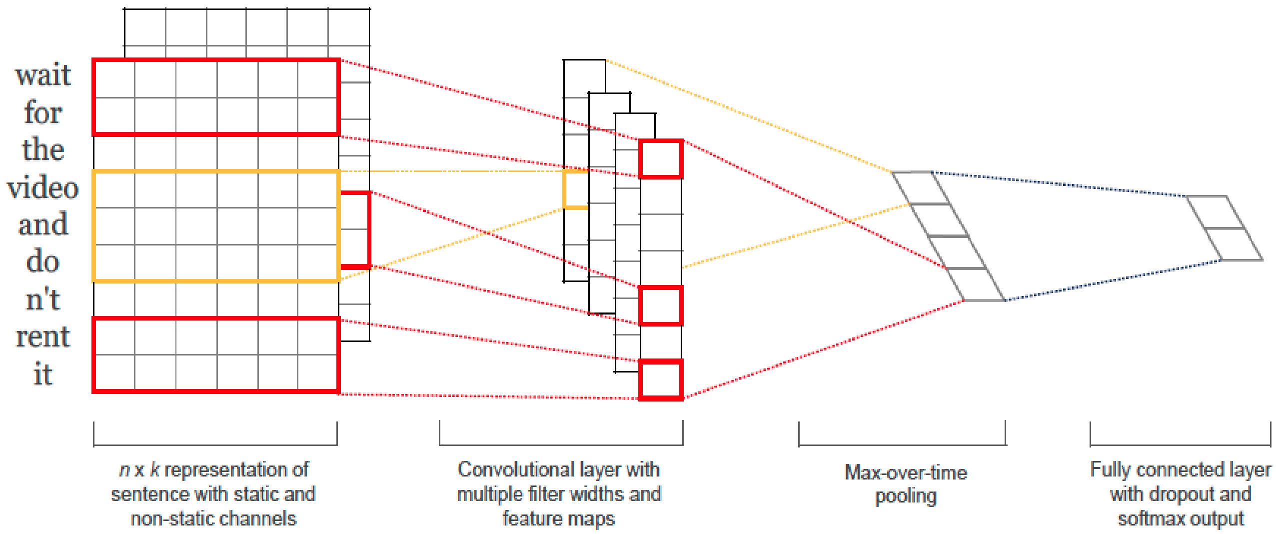
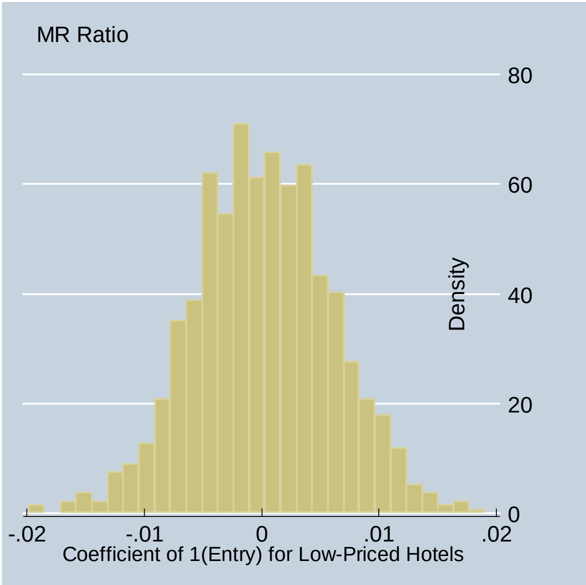


Figure W4.3. A Placebo Test of Randomized Assignment of Home Sharing's Entry

(a) Lower-priced Hotels




(b) Higher-priced Hotels