TWO STUDIES ON USER-GENERATED CONTENT IN ONLINE PLATFORMS: REVIEW VALENCE, SELF-PRESENTATION, AND SALES PERFORMANCE

LIU FUZHEN

MPhil

The Hong Kong Polytechnic University

2021

The Hong Kong Polytechnic University

Department of Logistics and Maritime Studies

Two Studies on User-generated Content in Online Platforms: Review Valence, Self-presentation, and Sales Performance

LIU Fuzhen

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Philosophy

May 2021

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

_________________________________ (Signed)

______________LIU Fuzhen_________ (Name of student)

Abstract

When the sharing economy encounters explosive growth of user-generated content (UGC), information asymmetry between sellers and buyers is a salient concern in carrying out online transactions. Trust-building signals gain prominence for customers to reduce their uncertainties in online activities. On the value of trust-building signals, it is desirable to understand how and why they matter for sales performance in the sharing economy context. So far, the literature remains unclear on two questions: 1) what signals from customer- and provider-generated content determine customer purchases? 2) what factors moderate the relationship between these trust-building signals and sales performance? In this thesis, we conduct two studies on trust-building signals derived from UGC in online platforms to identify and examine the factors influencing the signals-performance link from the contingency theory perspective.

In the first study, using data from Xiaozhu.com covering seventeen cities in China, we aim to examine the effects of review sentiment and average rating as well as their interactions with seller popularity and property quantity on sales performance. In terms of text mining, we adopt Naïve Bayes (NB) to extract review sentiment. Through a time-lagged regression model and city and genre fixed effects controlled for this model, we find no difference between review sentiment and average rating in positively influencing sales performance. Notably, seller popularity strengthens while property quantity weakens the positive impact of review valence on sales performance. More importantly, the positive effect of review valence on sales performance is more prominent for hosts characterized by popularity and personalization, and such hosts own fewer listings to achieve more historical sales. Our findings provide new insights by comparing the performance effects of review sentiment and average rating and uncovering the moderating role of seller popularity and property quantity. The implications are helpful for service providers to enhance their service operations management and for policy makers to better regulate the sharing economy.

In the second study, using data from Airbnb covering four cities in the United States, we investigate the effects of host self-presentation and social orientation embedded in self-presentation as well as their interactions with customer- (i.e., review rating) and marketer-generated (i.e., superhost) reputations on sales performance.

Latent Dirichlet allocation (LDA) is employed to identify the topics included in self-presentation, and we find two formats, including social-oriented and official-oriented topics. Furthermore, we use the support vector machine (SVM) algorithm to predict the topic of social-oriented self-presentation. Using econometric analysis, we find that self-presentation, especially for social-oriented self-presentation, positively influences sales performance. Moreover, we provide evidence supporting that customer-generated reputation (i.e., review rating) strengthens, but marketer-generated reputation (i.e., superhost) weakens the positive effects of self-presentation on sales performance. This research provides new insights by uncovering the complementary effect of customer-generated reputation and the suppression effect of marketer-generated reputation in the link between self-presentation and sales performance. Our findings offer managerial implications for service providers to formulate marketing strategies through building self-presentation and an online reputation.

Keywords: Review valence; Seller popularity; Property quantity; Sales performance; Self-presentation; Social orientation; Reputation; Sharing economy

Publications Arising from the Thesis

Published Journal Papers:

Liu, Fuzhen, Kee-Hung Lai, Jiang Wu, and Wenjing Duan. "Listening to online reviews: A mixed-methods investigation of customer experience in the sharing economy." *Decision Support Systems* (2021): 113609.

Papers Under Review:

Liu Fuzhen, Lai Kee-hung, Wu Jiang, Xin (Robert) Luo. Examining Electronic Word of Mouth in the Sharing Economy and the Role of Price and Responsiveness on Sales Performance: A Mixed-methods Investigation. *International Journal of Electronic Commerce* (Under first-round review)

Liu Fuzhen, Lai Kee-hung, He Chaocheng. A contingency perspective to investigate the effects of review valence on sales performance: a mixed-methods approach. *Journal of Business Research* (Under first-round review)

Liu Fuzhen, Lai Kee-hung. The value of self-presentation and online reputation in sharing economy operations. *Journal of Operations Management* (Under first-round review)

Acknowledgements

The journey at The Hong Kong Polytechnic University is full of wonderful memories that are worth my cherishing. Encounters and experiences during my two-year MPhil study are precious. I would like to express my sincere gratitude to all people who have provided me support and encouragement to fight for dreams.

To begin with, I am very grateful that my supervisor, Prof. Kee-hung LAI (Mike Lai), gives me patient and insightful guidance on how to conduct valuable research. For my research work, he always puts great effort into improving the readability of draft manuscripts with a rigorous academic attitude and dedication. His academic passion has deeply inspired me, and I can stay optimistic and curious to conduct academic research no matter what difficulties I encounter. More importantly, to keep healthy and promote team cohesion, he spares his time to play sports and go hiking with us like a friend. Again, I express my greatest gratitude to my supervisor for his selfless help.

Second, I would like to thank the Board of Examiners: Prof. Wai Ting Eric NGAI, Dr. Michael C.L. CHAU, and Dr. XU Jingjun David. I sincerely appreciate that they participate in my oral defense and give me constructive comments on improving the quality of my thesis.

Third, I would like to express my sincere and special gratitude to all the academic staff, research students, and administrative staff in the Department of Logistics and Maritime Studies who get along with me in these two years.

Last but not least, I sincerely appreciate that I have my beloved family and friends, who firmly support my choices and accompany me forever.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Research Background and Research Questions

With the advancement of information and communication technology (ICT), the sharing economy has piqued mounting attention in recent years (Zamani et al., 2019; Xu, 2020). As a representative of the leading accommodation-sharing economy, Airbnb has owned over 4 million hosts, who have entertained more than 800 million guests. It covers 100,000 cities in approximately 220 countries and regions[1]. Although the sharing economy has achieved rapid explosion at an unprecedented rate, trust is of paramount concern that hinders customer choices because of the lack of seller-buyer interactions (Zamani et al., 2019). The information asymmetry arising from limited interactions before purchase makes customers feel safety risks and privacy concerns in the sharing economy (Ert et al., 2016). According to a report, 52% of survey respondents think that personal safety is their greatest concern, and 58% of consumers in the United Kingdom and the United States point out that the risks outweigh the benefits in choosing the sharing economy[2]. To mitigate risks from purchasing an unknown product, customers resort to multiple signals displayed on online platforms.

The sharing economy heavily relies on signals in online platforms to enhance the seller-buyer trust before purchases (Han et al., 2019). As a triadic business model, peer-to-peer (P2P) accommodations consist of three parties: service providers, service consumers, and service enablers (e.g., the platform) (Hua et al., 2020). From this standpoint, service providers and consumers are the users of computer-mediated platforms. In the online environment, user-generated content (UGC) is crucial in sustaining the platform-mediated economy (Li et al., 2019). It conveys multiple and diverse signals that are helpful for service providers to implement marketing strategies and for customers to support their purchase decisions (Wang et al., 2016). Following previous research (Biswas et al., 2020), customer-generated content (CGC) and marketer-generated content (MGC) constitute UGC. In the sharing economy, UGC is formed by the information disclosed from providers, platforms, and consumers, determining customer purchase behavior (Xu et al., 2021). As two pivotal elements of UGC, MGC and CGC are the primary forms of disclosing information in P2P accommodations, carrying signals that help customers make a value judgment on

[1] https://news.airbnb.com/about-us/ (accessed on April 30, 2021)
[2] https://www.lloyds.com/about-lloyds/media-centre/press-releases/massive-sharing-economy-potential-remains-unrealised (accessed on April 30, 2021)

unknown products (Liang et al., 2020). It has been reported that online customer reviews, host profiles, and listing descriptions are the crucial information sources that develop customer trust (Xu et al., 2021). Moreover, scholars have agreed that MGC and CGC significantly influence customer purchase decisions (Liang et al., 2020)

Despite the momentum regarding the literature on CGC and MGC in swaying customer purchases, the findings from previous research are mixed and contradictory. In terms of the role of CGC, scholars have not arrived at a consensus over the impact of review valence on customer purchase in the sharing economy. Some scholars have supported that review valence positively influences price (Ren et al., 2021) and purchase intentions (Liang et al., 2017). However, other researchers have argued that review volume exerts a more decisive influence on purchasing P2P accommodations than review valence (Fu et al., 2021). In addition, evidence has shown that review scores have no significant impact on listing price or booking intentions (Ert et al., 2016). One possible reason is that compared to the traditional economy, the sharing economy may receive a relatively high review score (Bulchand-Gidumal & Melian-Gonzalez, 2020). Only reading peer customer reviews is risky for choosing P2P accommodations where consumers interact with service providers before, during, and after their stay. Customers may seek other signals regarding service or product quality to ensure that service providers are approachable and trustworthy (Yao et al., 2019). Prior research has shown that provider-specific features implied in host profiles are pivotal for customers to judge service quality and make purchase decisions (Wu et al., 2017). It is worthwhile to conduct an in-depth investigation on examining the role of review valence in influencing sales performance and gauge provider-specific factors affecting the relationship between review valence and sales performance.

Moreover, as a manifestation of MGC, self-presentation is a self-marketing way to advertise products and form a virtual reputation for service providers (Pera et al., 2016). It can bring trust perception (Zhang et al., 2020) and booking intentions (Tussyadiah & Park, 2018). However, scholars have mentioned no studies showing a significant relationship between provider self-presentation and sales performance (Xu et al., 2021). In this thesis, we aim to investigate the impact of self-presentation on sales performance. In addition to self-presentation acting a role in building an online reputation, there exist two kinds of traditional reputations in the accommodation-sharing economy (Abrate & Viglia, 2019). Online reputation is crucial in P2P accommodations because it helps create the brand for product and gain customer trust

(Mauri et al., 2018; Abrate & Viglia, 2019; Zervas et al., 2020). Prior research has suggested that various reputation-related signals affect each other and further influence sales performance (Wang et al., 2016). Nevertheless, there is no evidence supporting whether reputation moderates the impact of self-presentation on sales performance. It calls for an attempt to examine the effect of self-presentation and its interactions with reputation mechanisms on sales performance in the sharing economy.

Accordingly, from the perspectives of signal-based decisions and marketing strategies, this thesis targets to examine the effects of review valence (CGC) and provider self-presentation (MGC) on sales performance. In addition, we are motivated to investigate the boundary condition of the relationship between UGC (i.e., MGC and CGC) and sales performance. Thus, the following research questions guide the development of this thesis:

RQs of Study 1: From the perspective of signal receivers, does review valence determine sales performance? If so, whether provider-specific factors embedded in host profiles may moderate the impact of review valence on sales performance?

RQs of Study 2: From the perspective of signal senders, does self-presentation promote sales performance? If so, whether reputation mechanisms can moderate the effect of self-presentation on sales performance?

1.2 Motivations for Studying UGC

1.2.1 Motivations for investigating the effects of CGC on sales performance

As the primary element of CGC, online customer reviews have attracted scholarly attention in studying customer behavior (Zhao et al., 2015; Wang et al., 2016). Most studies have starkly focused on gauging customer experience extracted from textual reviews or investigating the outcomes of numerical reviews (Thomsen & Jeong, 2020; Ren et al., 2021). However, less attention is paid to link textual reviews and numerical reviews to customer purchase behavior. To have an in-depth understanding on the role of CGC, we consider simultaneously textual reviews (i.e., review sentiment) and numerical reviews (i.e., average rating) and examine their effects on sales performance.

Despite the momentum of the literature on examining the consequences of online customer reviews (Zhao et al., 2015; Wang et al., 2016), scant literature on gauging the contingency factors affecting the relationship between review valence and sales performance exists. In P2P accommodations, host-guest interaction is a salient feature before, during, and after the customer encounter experience (Wu et al., 2021). This unique feature of P2P accommodations may cause privacy concerns and security risks rising from service providers (Bulchand-Gidumal & Melian-Gonzalez, 2020). To develop trustworthiness toward service providers, host profiles are crucial as online customer reviews because they convey service attributes (Wu et al., 2017). Notably, property quantity is prominent in host profiles, which attracts customers' attention and affects their judgment over product and service quality (Xie et al., 2019). First, compared to hotel managers, property hosts lack the experience of managing the consumer-brand relationship and recovering service failures in responding to customer reviews (Liang et al., 2021). For instance, hosts in P2P accommodations tend to reply more to positive reviews than to negative reviews and even negatively respond to negative reviews. Second, unlike professional service providers, property hosts perform informally during the encounter experience, making customers feel psychologically close to hosts (Ju et al., 2019). For example, customers often build an intimate relationship with service providers because informal communication makes guests feel the warm hospitableness from hosts. Third, compared to professional hosts who own multiple listings, personal hosts have more energy and time assigned to each listing to offer high-quality services (Xie & Mao, 2017). Like property quantity, seller popularity (i.e., historical sales volume) is a quantity-related signals embedded in host profiles that capture customers' attention. Notably, seller popularity may trigger herd behavior because it reflects collective choices (Cheung et al., 2014; Xie et al., 2019). Evidence has shown that seller popularity can moderate the effect of online customer reviews on customer purchases (Xie et al., 2019). Drawing on the signaling theory and observational learning theory, seller popularity and online customer reviews are the signals of depicting the popularity, which can increase product awareness and influence customer purchases (Xie et al., 2019). Based on the above, we aim to investigate the moderating role of seller popularity and property quantity on the relationship between review valence and sales performance.

1.2.2 Motivations for examining the effects of MGC on sales performance

Although scholars have found the crucial role of MGC in determining customer behavior, the literature on investigating provider self-presentation is scattered (Xu et al., 2021). As an effective way of self-marketing, information disclosure has gained mounting popularity from scholars and practitioners (Xie et al., 2019). In P2P accommodations, self-presentation is a manifestation of information disclosure from service providers. Prior studies have examined the outcomes of information disclosure on customer satisfaction (Moon et al., 2019), trust perception (Zhang et al., 2018), and customer purchases (Tussyadiah & Park, 2018). In particular, scholars have found that the linguistic and semantic features of self-presentation can affect trust formation and booking intentions (Zhang et al., 2020). Yet, less attention is paid to linking provider self-presentation and its semantic feature to actual customer purchases. Scholars have reported a lack of substantial evidence to support the significant link between provider self-presentation and consumer purchases (Xu et al., 2021). To better understand the role of self-presentation, we consider self-presentation and its semantic feature as the antecedents of sales performance in the accommodation-sharing economy.

The topics over what contextual factors affect the relationship between self-presentation and sales performance are under-investigated. Unlike hotels normally endowed by brand chains, P2P accommodation is a business model with limited brand support (Fu et al., 2021). Like online reputation, self-presentation is a powerful self-storytelling to build a personal and digital reputation (Pera et al., 2016). Under the information asymmetry, online reputation is of paramount importance in swaying customer decisions because it signals the virtual brand for unknown service or product (Mauri et al., 2018). The development of strong online reputation mechanisms has enabled the sharing economy (Hamari et al., 2016) and facilitated the dynamic evolution of trust (Ert & Fleischer, 2019). Specifically, online reputation in P2P accommodations manifests in online customer reviews, Zhima credit, and the badge of "superhost". As a trust-building signal, online reputation can moderate the impact of visual clues on customer purchases (Abrate & Viglia, 2019). Considering that self-presentation can build a virtual reputation for service providers, we aim to investigate whether two kinds of traditional reputations, such as customer- (i.e., review rating) and marketer-generated reputations (i.e., the badge of "superhost"), moderate the impact of self-presentation on sales performance.

1.3 Research Objectives

To address the above research questions, we aim to examine the effects of UGC (i.e., CGC and MGC) on sales performance as well as the contingency factors (i.e., provider-specific features and traditional reputations) in influencing these relationships. Thus, this thesis has the following objectives to guide our investigation.

(1) Provide empirical evidence over the relationship between review valence (i.e., review sentiment and average rating) and sales performance in P2P accommodations.

(2) Examine whether provider-specific factors (i.e., seller popularity and property quantity) moderate the relationship between review valence and sales performance.

(3) Identify the semantic features of provider self-presentation.

(4) Investigate the impact of self-presentation as well as social-oriented self-presentation on sales performance.

(5) Explore whether the effects of self-presentation on sales performance depend on two kinds of traditional reputations (i.e., customer- and marketer-generated reputations).

(6) Offer theoretical and managerial suggestions for improving service operations management and formulating effective marketing strategies in the sharing economy as well as service businesses.

1.4 Research Framework and Methods

To deeply and systematically gauge the role of UGC in supporting customer decisions and marketing the products, we conduct two interrelated studies to examine the effects of CGC and MGC on sales performance, respectively. Study 1 investigates and compares the impact of review sentiment and average rating on sales performance as well as the moderating effects of provider-specific features (i.e., property quantity and seller popularity). In online platforms, CGC and MGC are two essential aspects of UGC. Indeed, online customer reviews as the manifestation of CGC determine customer purchases (Fu et al., 2021), provider self-presentation as the form of MGC is also crucial in influencing customer decisions (García et al., 2019). However, the scant literature on UGC emphasizes the importance of MGC in the accommodation-sharing economy (Liang et al., 2020). Study 2 examines the impact of self-presentation and its semantic feature on sales performance and tests the moderating role of two kinds of traditional reputations (i.e., customer- and marketer-generated reputations) on the above relationships. It can extend the existing literature proposing no significant relationship between self-presentation and sales performance (Xu et al., 2021). In summary, this thesis consists of two perspectives from customers and service providers to examine the critical role of CGC and MGC in influencing sales performance.

In terms of research theories, Study 1 uses the signaling theory and observational learning theory to illustrate the effects of review valence and its interactions with provider-specific features on sales performance. These theories help explain customer behavior from the perspective of information processing and highlight the paramount importance of visual signals in the context of information asymmetry (Xie et al., 2019). Study 2 adopts the social distance theory and signaling theory to examine the influence of self-presentation and its interactions with online reputation on sales performance.

Unlike Study 1 focusing on information processing when customers make a purchase decision, Study 2 introduces the social distance theory because self-presentation is a way of host-guest interaction and makes customers feel psychologically close to service providers (Moon et al., 2019). Moreover, Study 1 and Study 2 employ a mixed-methods approach of text mining and econometric analysis because we link unstructured text and structured data to sales performance. Due to the popularity of the accommodation-sharing economy (Pitt et al., 2021), we focus on P2P accommodations to pursue our research goals. Specifically, Study 1 observes the Chinese guests from Xiaozhu.com because numerous studies have investigated customer experience from online customer reviews in Airbnb with less attention paid to Chinese P2P accommodations (Luo & Tang, 2019; Thomsen & Jeong, 2020; Ren et al., 2021). In addition, considering the availability of self-presentation in the form of unstructured text in Airbnb, Study 2 targes at Airbnb users in the United States. Figure 1.1 presents the research framework that guides this thesis.

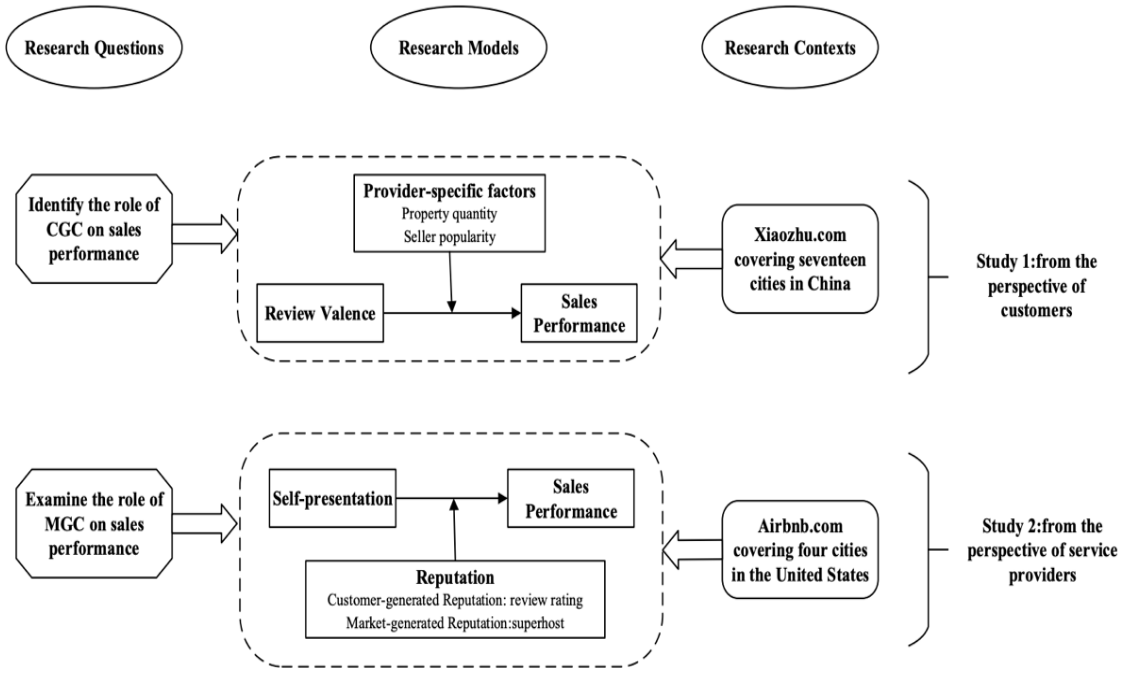


Figure 1.1 Overall research framework for this thesis

1.5 Research Significance

The thesis contributes new knowledge to the existing literature in several aspects. To clarify the research significance, we summarize the theoretical contribution of each study, as follows.

Study 1 advances the knowledge on online customer reviews by demonstrating the positive effects of review valence on sales performance in P2P accommodations. It can corroborate existing research on the positive link between review valence and sales performance in a service industry (Mariani & Borghi, 2020) and reconcile the debate over the role of review valence in the accommodation-sharing economy (Bulchand-Gidumal & Melian-Gonzalez, 2020). In addition, we also uncover that there is no substantial difference between review sentiment and average rating in illustrating sales performance. This finding can extend existing research just focusing on the relationship between review valence and sales performance, instead of comparing the performance effects caused by different dimensions of review valence (Choi et al., 2018). More importantly, Study 1 offers new insights on the moderating role of provider attributes in the sharing economy by revealing the complementary effect of seller popularity and the suppression effect of property quantity on the relationship between review valence and sales performance. These findings confirm that host attributes are crucial as listing characteristics and product quality in the service businesses (Wu et al., 2017; Xie et al., 2019). Finally, our findings extend the signaling theory and observational learning theory, indicating that signals will interact with each other and compete for customers' attentions (Xie et al., 2019).

Study 2 deepens the understanding of information disclosure by identifying the semantic features of self-presentation and demonstrating the positive effects of self-presentation on sales performance. It confirms that self-presentation is a way of reducing social distance because it depicts host-guest interaction and builds a virtual reputation for service providers (Pera et al., 2016; García et al., 2019). In addition, Study 2 extends the literature on online reputation by uncovering the complementary effect of customer-generated reputation and the substitution effect of marketer-generated reputation in the link between self-presentation and sales performance. It not only verifies that online reputation is crucial for forming a virtual self to reduce social distance in the accommodation-sharing economy (Mauri et al., 2018; Abrate & Viglia, 2019), but also further indicates that diverse reputation-related signals can interact and work together to determine customer purchase (Wang et al., 2016). Notably, our findings extend the signaling theory by revealing that trust-building signals may complement (substitute) each other when they are from diverse (same) sources.

1.6 Structure of the Dissertation

This dissertation includes five chapters to deeply and systematically illustrate the importance of UGC in determining sales performance, as follows.

Chapter 1 illustrates our research background by proposing research questions. It depicts the motivations of this thesis, research objectives, and research significance. In addition, it also outlines the overall research framework and introduces the structure of this dissertation.

Chapter 2 elaborates on the literature review laying a theoretical background for identifying research gaps implied in the literature and helping readers to understand this thesis. In particular, we summarize the factors influencing sales performance in P2P accommodations from two perspectives of MGC and CGC.

Chapter 3 sheds light on the effects of review valence (i.e., review sentiment and average rating) and its interactions with seller popularity and property quantity on sales performance using a mixed-methods approach of text mining and econometric analysis. Based on the dataset from Xiaozhu.com, we capture key findings and offer implications for academic researchers and practitioners. Moreover, the limitations are discussed, which are potential areas to be addressed in future research.

Chapter 4 explains the impact of self-presentation as well as its interactions with customer- and marketer-generated reputations on sales performance using a mixed-methods approach of latent Dirichlet allocation (LDA) and econometric method. Using the dataset from Airbnb, we draw some conclusions and extend the existing literature. Besides, we also give some theoretical and managerial implications. In addition, this Chapter proposes some limitations that motivate future research.

Chapter 5 summarizes the key findings of this thesis, concludes the implications in theoretical and managerial aspects, and outlines future research directions to address the potential limitations in this thesis.

Chapter 2: Literature Review

2.1 Sharing Economy

Sharing economy refers to the utilization of idle resources or services temporarily transferred from owners to customers (Belk, 2014). As a kind of experiential service industry, sharing economy is operated on the basis of trust with the help of computer-mediated platforms (Ert et al., 2016). Unlike traditional e-commerce, the sharing economy allows service providers and customers to build face-to-face interactions and communications throughout the guest experience after online booking (Wu et al., 2021). In the presence of information asymmetry, customers perceive multiple risks pertaining to security and privacy in receiving uncertain service or product quality. One way for them to mitigate these risks is to seek trust-building signals in online platforms (Xie et al., 2019). They do so because these signals are essential in helping judge product or service quality and forming customer trust, especially for home-sharing services (Xie & Mao, 2017). In the sharing economy, trust in service providers is a crucial element in promoting customer decisions (Wu et al., 2017).

Peer-to-peer (P2P) accommodation, as a home-sharing service, has been widely selected as the observation sample of the sharing economy to examine its service operations management (Ju et al., 2019; Liang et al., 2020; Wu et al., 2021). There are two categories of home-sharing services: monetized (e.g., Airbnb) and non-monetized (e.g., Couchsurfing). The former provides services for profit, while the latter emphasizes community engagement based on trust and reciprocity. The monetized home-sharing has achieved explosive growth globally, attracting scholarly attention (Liang et al., 2020; Wu et al., 2021). In P2P accommodations, hospitality and sociability are two major elements that customers care about. The demand for service quality is rising due to financial involvement in the exchange of monetized sharing services (Ikkala et al., 2015). However, customers encounter more perceived risks from privacy and security (Ert et al., 2016). The trust-building signals in online platforms are crucial for customers to make a proper judgment over service or product quality (Xie et al., 2019).

2.2 CGC in the Accommodation-sharing Economy

P2P accommodation is a business model of experiential consumption, heavily relying on electronic word-of-mouth to create a virtual brand (Han et al., 2019). As an essential kind of CGC, online customer review has risen to prominence as the reflection of customer experience and opinions (Fu et al., 2021). In e-commerce, where information asymmetry is apparent, customers tend to trust collective wisdom and follow peer opinions to reduce perceived risks (Cheung et al., 2014). Scholars have provided strong evidence supporting that online customer reviews exert a significant influence on customer trust and behavior (Xie et al., 2019; Fu et al., 2021). For example, review volume in the accommodation sector signals popularity cues of products, which determines customer purchase decisions (Fu et al., 2021).

Table 2.1 summarizes prior studies on investigating online customer reviews in the accommodation-sharing economy. The first body of literature has focused on numerical reviews, such as review valence (e.g., average ratings) and volume (i.e., the number of reviews). Numerous studies have demonstrated that average rating and review number positively influence price (Ren et al., 2021), booking intentions (Fu et al., 2021), and reputation (Liang et al., 2017). Referring to previous research (Duan et al., 2008), review valence plays a persuasive role in determining customer decisions, and review volume helps customers be aware of the product. The second block of the previous literature has investigated customer experience extracted from textual reviews (Cheng & Jin, 2019; Cheng et al., 2019; Thomsen & Jeong, 2020). It not only examines what attributes of P2P accommodations that customers care most about but also helps understand what sentiment is expressed in textual reviews (Luo & Tang, 2019). In particular, it has been reported that textual reviews are significantly associated with numerical ratings, which can illustrate customer satisfaction and trust perception (Luo & Tang, 2019; Pitt et al., 2021). The last block of existing research has linked textual reviews to numerical reviews in P2P accommodations (Biswas et al., 2020).

Overall, the extant literature has mainly focused on Airbnb guests to investigate the outcomes of number-specific review characteristics, customer experience extracted from text-specific review characteristics, and the relationship between textual reviews and numerical reviews (Luo & Tang, 2019; Thomsen & Jeong, 2020; Ren et al., 2021). However, less attention is paid to examining review valence by simultaneously considering numerical reviews and textual reviews in the Chinese accommodation-sharing economy. In P2P accommodations, service provider plays a crucial role in influencing customer decisions (Wu et al., 2017). Yet, the topic over whether provider-specific features as the contingency factors moderate the relationship between review valence and sales performance is under-investigated.

Table 2.1 Previous studies on online customer reviews

Content	Sources	Key findings	Samples	Theories
The outcomes of numerical review	(Ren et al., 2021)	Online review valance (average star ratings) and volume (number of reviews) positively influence the price	Airbnb	Attribute substitution theory
	(Fu et al., 2021)	Review volume exerts a more decisive influence on booking intentions in P2P accommodations than review valence does.	Not specific	Signaling theory
	(Ert et al., 2016)	Review valence (i.e., review scores) does not significantly affect listing price or booking intentions	Airbnb	Trust theory
	(Liang et al., 2017)	Review volume and review valence are positively associated with the "superhost" badge in P2P accommodations	Airbnb	rational action theory
Customer experience extracted from textual review	(Thomsen & Jeong, 2020)	Key themes extracted from online customer review include listing specifics, host attributes, and recommendation. Specifically, the score of "cleanliness" positively affects the overall review rating	Airbnb	Not specific
	(Cheng et al., 2019)	Listing characteristics, host attributes, and description and evaluation are the salient cognitive themes. In addition, several themes (host attributes, room aesthetics and location, and room description) positively affect trust (benevolence, ability, and integrity). Location and host attributes determine overall trust perception.	Airbnb	Trust theory
	(Luo & Tang, 2019)	Five aspects of textual reviews include host communication, experience, house location, service/product, and value. In addition, aspect-specific sentiment positively influences the satisfaction (i.e., overall numerical rating)	Airbnb	Not specific
	(Pitt et al., 2021)	Textual reviews are significantly associated with numerical rating.	Not specific	Not specific
	(Xu, 2020)	Social interaction and economic value are attached more importance for customers at a higher sharing-level accommodation.	Airbnb	social penetration and social exchange theories
The relationship between textual and numerical reviews	(Cheng & Jin, 2019)	Three key attributes are "location", "amenities", and "host".	Airbnb	Not specific
	(Biswas et al., 2020)	Numerical rating (i.e., overall review scores) and review sentiment (i.e., negative sentiments) significantly affect review volume (i.e., the number of reviews)	Airbnb	Not specific

2.3 MGC in the Accommodation-sharing Economy

As a manifestation of MGC, provider self-presentation is as crucial as listing specifics in determining customers' decision-making (Xu et al., 2021). Especially in P2P accommodations, where service providers participate in customer experience, the platform empowers service providers to facilitate information exchange by disclosing themselves. Previous research has demonstrated that host attributes embedded in MGC have a significant effect on customer trust (Wu et al., 2017). Indeed, self-presentation has become a self-marketing strategy for service providers to advertise products and attract customers (Zhang et al., 2020). It helps outline a virtual image of service providers and form customer trust towards service providers (Tussyadiah & Park, 2018).

Table 2.2 epitomizes a cluster of literature on information disclosure from service providers in the accommodation-sharing economy. The first body of prior research has focused on investigating the impact of information disclosure on customer satisfaction and loyalty (Moon et al., 2019) and customer purchase (Xie et al., 2019; Xu et al., 2021). The second block of the previous literature has examined the linguistic and semantic features of self-presentation in influencing trust perception (Zhang et al., 2018; Zhang et al., 2020), booking intentions (Tussyadiah & Park, 2018), and sellers' revenues (García et al., 2019). Specifically, the readability, the sentiment, and perspective taking of self-presentation significantly affect trust formation (Zhang et al., 2020). Moreover, self-presentation (i.e., social orientation and travel amateur) exerts a significant influence on customer purchases (Tussyadiah & Park, 2018; García et al., 2019).

Overall, despite that numerous studies have emphasized the role of information disclosure in the accommodation-sharing economy, these studies have starkly focused on trust theory (Tussyadiah & Park, 2018; Zhang et al., 2018). However, the literature using construal-level theory and signaling theory to illustrate self-presentation is sparse. Moreover, existing research lacks substantial evidence to link provider self-presentation to sales performance as well as examine the contingency factors that affect the performance effects of self-presentation (Xu et al., 2021).

Table 2.2 Previous studies on provider descriptions

Research Topics	Sources	Key findings	Samples	Theories
The impact of information disclosure on customer attitude and behavior	(Moon et al., 2019)	Self-disclosure is significantly associated with encounter satisfaction, WOM intention, and continuous intention to use	Not specific	Social penetration theory
	(Xie et al., 2019)	Sales history disclosure positively influences customer purchases.	Xiaozhu	observational learning theory
	(Xu et al., 2021)	Information disclosure from providers (i.e., the amount of social identity, verification, the responsiveness, and the disclosure of their social network profiles) significantly influence customer purchase behavior	Airbnb	signaling theory, social exchange theory, transaction cost theory, and information process theory
The outcomes of the linguistic and semantic features	(Zhang et al., 2020)	The linguistic (i.e., readability, sentiment intensity, perspective taking) and semantic features (i.e.,	Airbnb	Trust theory

embedded in self-presentation		talking more about family relationship, openness, service, and travel experience) of self-description significantly affect trust perception.		
	(Tussyadiah & Park, 2018)	Hosts present themselves as a well-traveled individual or an individual of a certain profession. The well-traveled hosts embedded in self-presentation gain higher levels of perceived trust and more booking intentions.	Airbnb	Trust theory
	(García et al., 2019)	The social-oriented self-description and the self-presentation length significantly affect sellers' revenues	Airbnb	construal-level theory
	(Zhang et al., 2018)	The positive sentiment in self-presentation or profile photo is essential in trust formation	Airbnb	Trust theory

Chapter 3 Study 1: Review Valence and Sales Performance: The Role of Seller Popularity and Property Quantity

3.1 Introduction

With technological advancement in the industry 4.0 era, the lack of trust is still conspicuous in the platform-mediated economy due to limited information access (Ert et al., 2016). To make a proper purchase decision, customers are likely to resort to online customer reviews because they trust peer evaluations (Jia & Liu, 2018). According to Stackla[3], customers have said that user-generated content (UGC) is authentic and impactful compared to brand-created content and influencer-created content. In addition, seventy-nine percent of customers have reported that UGC highly affects their purchasing decisions. Indeed, electronic word-of-mouth (e-WOM) has gradually become an advertising strategy to obtain a competitive advantage in attracting potential customers (Raguseo & Vitari, 2017). Unlike multinational companies investing millions of dollars in the traditional advertisement, Starbucks relies on e-WOM as a marketing strategy to promote products and acquire customers[4]. Starbucks set up a website called “My Starbucks Idea” to encourage customers to share opinions and give feedback to enhance the service. Moreover, due to its superior product quality and store experience (i.e., free high-speed Internet), customers spread their opinions favorably for Starbucks on famous social media platforms, such as Facebook, Twitter, Instagram, and YouTube. The widespread of e-WOM has benefited Starbucks for more sales because many people are aware of their products.

The sharing economy is community-based and heavily relies on e-WOM marketing because service providers participate in customer experience (Han et al., 2019). The sharing economy refers to the operations that service providers share their idle capacities and resources with strangers for offline experiential services after online booking (Ert et al., 2016). With the help of the platform as an enabler of information-carrying, the sharing economy has experienced explosive growth in recent years (Lim et al., 2021; Wu et al., 2021). According to the PWC report, sales revenue arising from the sharing economy is expected to reach 335 billion dollars in 2025, accounting for half of the rental market[5]. However, information asymmetry and the lack of interaction

[3] https://stackla.com/resources/reports/bridging-the-gap-consumer-marketing-perspectives-on-content-in-the-digital-age/ (accessed on April 30, 2021)
[4] http://www.brandba.se/blog/ewomcharacteristics (accessed on April 30, 2021)
[5] https://www.pwc.com/hu/en/kiadvanyok/assets/pdf/sharing-economy-en.pdf (accessed on April 30, 2021)

are two inhibitory factors hindering trust formation (Zamani et al., 2019). As a trust-building signal, e-WOM is crucial for consumers to mitigate risks and determine whether to purchase in the experiential service industry (Duan et al., 2008). Such a signal is helpful for service providers to obtain benefits by attracting customers because positive e-WOM significantly affects financial performance (Raguseo & Vitari, 2017), firm reputation (Gunter, 2018), booking intentions (Wang et al., 2016), and customer trust and repurchase intentions (Bulut & Karabulut, 2018).

Prior studies have investigated the consequences of e-WOM in the platform-mediated economy (Bulut & Karabulut, 2018). Nevertheless, a large body of research has starkly focused on e-WOM volume with inadequate attention paid to e-WOM valence (Cheung et al., 2014; Xie et al., 2019). Moreover, the performance effect of review valence is inconclusive with contradictory results (Duan et al., 2008; Yi et al., 2021). Some scholars have supported that review valence can promote customer engagement behaviors (Ye et al., 2019; Yi et al., 2021). For example, it has been found that positive review rating and review sentiment in hotels are favorable for Airbnb popularity (Yi et al., 2021). However, other scholars have indicated that review valence is irrelevant to customer purchase because it cannot truly reflect customer experience (Duan et al., 2008). These inconsistent findings are ambiguous for service providers to understand the influence of review valence on sales performance in the sharing economy. Against this background, we model review sentiment and average rating as the indexes of review valence to examine and compare their effects on sales performance in the accommodation-sharing economy.

Specifically, under what circumstances will review valence bring greater benefits on sales performance? In the accommodation-sharing economy, although e-WOM plays a persuasive role in swaying customer decisions, host profile is an attraction for customers when they search for product-related information online (Wu et al., 2017). Before, during, and after the stay, service providers participate in delivering customer experience. Specifically, host-guest interaction is an exceptional hospitality service for the sharing economy in the accommodation sector (Moon et al., 2019). It is not just offering a prescribed space to share with consumers but also providing extra assistance, such as inviting customers to a meal. As the quantity-related signals with awareness effect, property quantity and seller popularity are two salient tags embedded in host profile because they are external clues of depicting service quality (Duan et al., 2009; Xie & Mao, 2017). Property quantity refers to the number of listings managed by a host, reflecting whether a host is a personal or professional operator (Liang et al., 2020). Professional hosts include those service providers offering multiple listings. In contrast, personal hosts tend to operate with fewer or even single listings. There is a substantial difference in delivering quality service and pricing strategies between personal and professional hosts, which leads to their distinctive performance in operational and financial aspects (Li et al., 2016; Xie & Mao, 2019). Notably, the debate over the performance effect of property quantity is inconclusive with two rival opinions (Wu et al., 2017; Xie et al., 2019). Some scholars have supported that property quantity positively affects reservations because it shows hosts' ability and experience to provide adequate rooms for guests (Wu et al., 2017). Other researchers have found a negative effect of multi-listing ownership on customer purchase because the allocation of resources or service efforts by a multi-listing host to each listing decreases when the number of listings managed by a host increases (Xie & Mao, 2017). Property quantity is a double-edged sword in managing such service operations because balancing the trade-off between quality and quantity is common in economic decisions (Ellway, 2014). Similarly, the moderating role of property quantity in the relationship between trust-building signals and sales performance has negative and positive effects (Xie et al., 2019; Liang et al., 2020). These inconsistent results motivate us to further investigate on the role of property quantity in the accommodation-sharing economy.

Moreover, seller popularity is the index depicting the number of historical reservations a seller has achieved, reflecting peer customer purchases for a specific service provider (Chen, 2008; Duan et al., 2009). In e-commerce, where information asymmetry is frequent, customers resort to collective choices for reducing their perceived risks (Chen, 2008). Seller popularity is a signal generated by peer customers, reflecting collective choice and determining customer purchase behavior (Duan et al., 2009). In addition, seller popularity is as crucial as customer review because they reflect peer customer purchase and opinions, nurturing herd behavior to imitate others' behaviors for the purpose of lessening the uncertainties (Cheung et al., 2014). As the essential quantity-related signals embedded in host profiles, both seller popularity and property quantity are found beneficial directly for sales performance and indirectly through strengthening the effects of customer- or marketer-generated information (Viglia et al., 2014; Liang et al., 2020). So far, not much is known whether seller popularity and property quantity strengthen or weaken the influence of review valence on sales performance. This study aims to investigate their moderating role in the link between review valence and sales performance in the accommodation-sharing economy, respectively and collectively.

Drawing on the signaling theory and observational learning theory, we examine the relationship between review valence and sales performance as well as the moderating role of property quantity and seller popularity on this link. The signaling theory argues that individuals judge product or service quality based on what they receive from senders, affecting their subsequent behavior under information asymmetry (Spence, 2002). As a signal with the persuasive effect, review valence reflects customer experience from peers who have experienced a product or service. Property quantity indirectly mirrors service attributes as it demonstrates the capacity of managing resources. Seller popularity conveys a trust-building signal to prospective customers because it represents collective choices. In addition, the observational learning theory suggests that individuals tend to imitate peers' behaviors and trust collective choices, which is called herd behavior (Banerjee, 1992). Review valence and seller popularity depicting collective opinions and choices are generated by previous customers, leading prospective customers to follow herd behavior.

This study makes knowledge contributions in four aspects. First, we confirm and extend the signaling theory by uncovering the interactions of signals from customer reviews and host profiles on customer purchases. In addition, we further enhance the observational learning theory by demonstrating the herd behavior triggered by peer customer opinions (i.e., review valence) and choices (i.e., seller popularity) in online platforms. Second, we provide empirical evidence on the relationship between review valence and sales performance. More importantly, we have demonstrated that review sentiment is as crucial as average rating in influencing sales performance. Third, we offer new insights into the boundary conditions of the relationship between review valence and sales performance by analyzing the moderating role of property quantity and seller popularity in this relationship. Finally, we employ a mixed-methods approach integrating machine learning-based text mining with econometric analysis to test our hypotheses. This is the first attempt to gauge customer experience in a Chinese accommodation-sharing platform using a machine learning approach and link unstructured text (i.e., review sentiment) as well as structured data (i.e., average rating) to sales performance.

3.2 Literature Review

3.2.1 Signaling theory

A signal depicting the clue conveyed by senders to receivers concerning product or service quality is vital for customers to tackle perceived risks from information asymmetry and make a purchase decision (Spence, 2002). The signaling theory has been widely adopted in the marketing and information system literature because it can better illustrate transaction-specific communication and interactions (Cheung et al., 2014; Xie et al., 2019). In online platforms, customers are likely to resort to trust-building signals generated by previous customers and marketers. These signals are helpful for them to make a judgment over uncertain service or product quality, influencing their purchase decisions (Xie et al., 2019).

There is intensive offline interaction between guests and hosts after online booking in the sharing economy because a host may share a room with guests. The face-to-face interaction between guests and hosts may cause privacy and security concerns. In this context, customers need to know the product or service providers by assessing adequate information to reduce potential risks and make a proper decision (Xie et al., 2019). In the accommodation-sharing economy, marketer-generated and customer-generated contents are critical signals sent by sellers or buyers to prospective customers (Liang et al., 2020). Specifically, online customer reviews and supplier profiles are crucial for customers to evaluate service attributes and product characteristics (Liang et al., 2020). This study draws on the signaling theory to explain the importance of online customer reviews and supplier profiles as reliable information sources of building trust for interpreting customer purchases.

3.2.2 Online customer reviews

With the rapid growth of social media, online customer reviews have become a prominent information source as the reflection of peer evaluation on consumption experience. As one of the primary forms of e-WOM, online customer reviews are the powerful forces to influence customer decision-making because customers trust peer opinions (Li et al., 2019). Compared to traditional advertisements, e-WOM can effectively disseminate the overall quality of service providers and products, contributing to firms' financial performance (Raguseo & Vitari, 2017). Prior research investigating e-WOM has mainly focused on two aspects, review valence and review volume (Duan et al., 2008; Li et al., 2019). Scholars have consistently supported that the number of reviews as the index of review volume ushers customer purchase decisions because it increases product awareness (Duan et al., 2008). So far, there is still no consensus regarding the effect of review valence on sales performance (Ert et al., 2016). As the symbols of review valence, average rating and review sentiment (i.e., positive or negative) have been demonstrated to significantly affect sales performance (Raguseo & Vitari, 2017). Specifically, textual review (i.e., overall sentiment) can directly affect product sales and indirectly through influencing numerical rating (i.e., star rating) (Li et al., 2019). Moreover, textual sentiment analysis can better articulate customer opinions than review rating (Chen et al., 2017). However, unlike review volume exerting an awareness effect on sales, review rating has no persuasive influence on revenue (Duan et al., 2008).

In P2P accommodations, average rating determines the badge of “superhost” and then influences customer purchase decisions (Gunter, 2018). Compared to the traditional economy, the sharing economy receives a relatively higher product rating in general (Bulchand-Gidumal & Melian-Gonzalez, 2020; Santos et al., 2020). As the manifestation of guest satisfaction, textual sentiment can illustrate rating scores (Zhu et al., 2020). However, there is scant literature on comparing the effects of textual sentiment and average rating in influencing sales performance. Accordingly, we consider review sentiment and average rating as the indexes of review valence to examine and compare their effects on sales performance.

3.2.3 Supplier profiles

With the proliferation of social media, customer-generated content (CGC) and marketer-generated content (MGC) are two prominent parts of platform signals for information presentation and exchange (Liang et al., 2020). Although online customer reviews have provided helpful information for customers’ judgment, supplier profiles also play an essential role in swaying customer purchase decisions (Wu et al., 2017; Xie et al., 2019). Supplier profiles can portray the image of service providers, presenting a vivid description of their characteristics and service attributes. It can convey multiple signals in the form of text and pictures, helping customers to form an online impression of service providers (Xie et al., 2019). Evidence has shown that service attributes perceived from supplier profiles can influence customer trust and purchase behavior (Liang et al., 2020).

In the accommodation-sharing economy, supplier profiles are crucial as marketer-generated content for determining customer purchases (Liang et al., 2020). There are two critical quantity-related features of depicting service quality embedded in supplier profiles: seller popularity and property quantity (Xie et al., 2019). As an essential signal of mirroring host attributes, seller popularity is represented by the historical sales volume that a seller achieves (Cheung et al., 2014). It has been demonstrated to signal product or service quality, boosting customers' preference (Viglia et al., 2014). In e-commerce, seller popularity may trigger herd behavior among customers because it can reflect peer customer choices. In addition, customers tend to follow collective choices in online purchasing when they are uncertain about product quality (Chen, 2008). On the other hand, property quantity has attracted sparkling interest among scholars due to the contradictory view about its role in influencing service quality (Xie et al., 2019; Liang et al., 2020). Property quantity in P2P accommodations depicts whether a host is a professional or personal operator according to the listing volume that a host owns. Some scholars have supported property quantity positively and significantly affects customer purchases because it shows a host's ability to provide adequate rooms (Wu et al., 2017). Other researchers have found that multi-listing owners may not achieve excellent sales performance due to their limited capacity for managing the listings (Xie & Mao, 2017). In this study, we consider seller popularity (i.e., sales volume that a host achieves) and property quantity (i.e., listings volume that a host possesses) as provider-specific features and examine their moderating role in influencing the effect of review valence on sales performance.

3.3 Hypotheses Development

3.3.1 Review valence and sales performance

With the flourishing growth of online interactions, customer-generated content has risen to prominence as customers put more value on peer evaluation (Han et al., 2019). As a pivotal signal generated by customers, online customer review has gained mounting popularity in recent years (Kasabov, 2016). Especially in the service industry, where service quality determines customer experience, customers resort to online customer reviews for their decision-making because they consider peer opinions (Mariani & Borghi, 2020). The importance of review features, including review volume, review valence, and review variance, in swaying customer purchases, has been noticed because their performance value is identified (Floyd et al., 2014; You et al., 2015; Mariani & Borghi, 2020). Notably, two meta-analysis studies have provided strong evidence supporting that review valence plays a more prominent influence than review volume in influencing sales performance because the former can better illustrate customer opinions (Floyd et al., 2014; You et al., 2015).

As a signal with persuasive effect, review valence has been demonstrated to positively influence customer purchases (Mariani & Borghi, 2020). Specifically, as an index of review valence, average online rating exerts a positive effect on financial performance because it reflects customer experience after purchasing (Mariani & Borghi, 2020). In addition, scholars have found that positive review sentiment perceived by customers is positively associated with diagnostic value, which can increase customers' adoption of online reviews and promote their purchases (Jia & Liu, 2018). Moreover, review valence is one of the primary extrinsic cues of quality, which positively affects sales performance (Choi et al., 2018). In general, there seems to be a consensus across studies that review valence promotes sales performance. Accordingly, we propose the following hypotheses that

Hypothesis 1a: Review sentiment is positively associated with sales performance.

Hypothesis 1b: Average rating is positively associated with sales performance.

In a situation of lacking trust, customers tend to read more reviews posted by peers and compare review scores across stores (Ert et al., 2016). As the indexes of review valence, review sentiment and average rating have been found to positively affect customer purchases (Floyd et al., 2014; You et al., 2015). Compared to star rating, review sentiment plays a more critical role in purchasing an accommodation-sharing service because customers are likely to read more reviews to reduce risks (Chang & Wang, 2018). Moreover, textual sentiment provides a better summary of customer opinions than rating scores, indicating that review sentiment is more suitable to predict product sales than online rating (Chen et al., 2017).

On the other hand, as a traditional quality signal, star rating does not work well in explaining customer booking intentions in the accommodation-sharing economy (Yao et al., 2019). One possible reason is that many properties receive a relatively high average customer rating in the accommodation sector (Bulchand-Gidumal & Melian-Gonzalez, 2020). In this context, it is risky for customers to read few reviews to make a purchase decision. Instead, their purchase decisions depend more on review sentiment (Chang & Wang, 2018). Considering that review sentiment can better illustrate peer opinions, we argue that compared to average rating, review sentiment exerts a more decisive influence on sales performance. Thus, we hypothesize that

Hypothesis 2: The positive effect of review sentiment on sales performance is stronger than that of average rating.

3.3.2 Interactions of review valence with seller popularity

In online platforms, customer-generated content and behavior from previous customers may significantly influence the behaviors of prospective customers (Cheung et al., 2014). Under the information asymmetry, customers tend to follow peer opinions and actions to reduce risks and uncertainties, leading to a herd effect (Chen, 2008). The herd effect evolves from the social learning theory of observational learning. Observational learning theory refers to the situation that individuals make their decisions by observing others' behaviors without communication (Bandura & McClelland, 1977). This theoretical concept has gradually appeared in the economics literature and is redefined as an information cascade: observational learning makes individuals evaluate information and imitate peers' behaviors, called the herd effect (Banerjee, 1992). An example of herd behavior is that customers are enthusiastic about buying products listed at higher sales ranking in Alibaba's online shopping carnival (Xu et al., 2017). As the manifestation of collective opinions and actions, review valence and seller popularity can stimulate herd behavior, swaying customer purchase decisions (Cheung et al., 2014).

Drawing upon the signaling theory, seller popularity and review valence are two trust-building signals of depicting service or product quality in online platforms (Cheung et al., 2014). Seller popularity is measured by the sales volume that a host has achieved, which helps customers form trust and promote their purchases (Choi et al., 2018). In addition, it can mirror the popularity of service providers because many peers choose them (Cheung et al., 2014). Like seller popularity, review valence illustrates customer evaluation. It conveys a signal of product popularity, determining the purchase decisions of prospective customers (Chen et al., 2017). In the accommodation-sharing business model, customers have the opportunity of face-to-face interactions with service providers after online booking. However, customers often feel uncertain about unknown products before purchasing a product, and multiple trust-building signals displayed on online platforms jointly determine customer behavior (Xie et al., 2019). Considering the herd behavior and the quality signal from seller popularity and review valence (Banerjee, 1992; Choi et al., 2018), we argue that these two signals may affect each other and positively influence customer purchases. Accordingly, we propose that

Hypothesis 3a: Seller popularity strengthens the positive effect of review sentiment on sales performance.

Hypothesis 3b: Seller popularity strengthens the positive effect of average rating on sales performance.

3.3.3 Interactions of review valence with property quantity

In the sharing economy, home-sharing providers consist of regular hosts who own few listings and professional hosts who own multiple listings (Li et al., 2016; Xie & Mao, 2019). There are substantial differences in the financial and operational performance between professional and regular hosts (Xie et al., 2019; Xie & Mao, 2019). In Tujia, hosts have been tagged as professional and individual hosts, visible by customers. Scholars have found that property quantity plays a vital role in influencing listing performance through forming customer trust in service providers (Xie & Mao, 2017). In addition, the performance effects of host quality decrease when the number of listings owned by a host increases (Xie & Mao, 2017). From this standpoint, a balance between host quality and listing quantity is evident in the service delivery (Ellway, 2014).

Compared with personal service, professional interaction is more standardized, formal, profit-driven, and task-oriented. However, its limited affective and impersonal communication will hinder guest trust and customer satisfaction (Roschk & Gelbrich, 2017). In contrast, personal services can help customers trust the host through a friendship-like exchange. Personal interaction with the host will make customers feel "stay at home" due to perceived sincerity in guest experience (Ju et al., 2019). Thus, customers prefer personal hosts to professional hosts, as trust is easily built through personal touch and informal communication between hosts and guests (Roschk & Gelbrich, 2017). Considering the existence of the "trade-off" tension between quantity and quality and "trusted hospitality" from personal hosts, we argue that personal hosts can bring positive trust-building signals of service quality.

In online platforms, customer reviews and host profiles are essential information sources from which customers can make a proper judgment over the product or service providers (Xie et al., 2019). Evidence has shown that review volume exerts a positive influence on customer purchases, while property quantity negatively affects customer purchases (Xie et al., 2019). Like review volume, review valence can positively affect customer purchases because it represents customer experience and service quality (Han et al., 2019; Biswas et al., 2020). In P2P accommodations, review valence is a critical element of online reviews and manifests in review sentiment and average rating (Biswas et al., 2020). It has captured customers' attention because peer evaluation conveys persuasive information about the product (Han et al., 2019; Bulchand-Gidumal & Melian-Gonzalez, 2020). In contrast, multi-listings hosts may signal negative clues about service quality due to the limitation of resource allocation (Xie & Mao, 2017). Based on the above, we argue that positive review valence for personal hosts exerts a stronger influence on sales performance than that for professional hosts. Thus, we posit that

Hypothesis 4a: Property quantity weakens the positive effect of review sentiment on sales performance.

Hypothesis 4b: Property quantity weakens the positive effect of average rating on sales performance.

3.3.4 Interactions of review valence with seller popularity and property quantity

In the accommodation-sharing economy, seller popularity and property quantity are observable from host profiles, signaling service quality (Xie & Mao, 2017; Xie et al., 2019). The former depicts the sales volume that a seller has achieved, and the latter represents the listing volume that a seller has possessed. On the one hand, seller popularity positively influences customer purchases because customers perceive high-quality service or product from popular sellers (Cheung et al., 2014). On the other hand, the performance effects of host quality diminish when the number of listings owned by a host increases (Xie & Mao, 2017). It indicates a negative relationship between property quantity (i.e., the number of listings) and sales performance, indicating that a listing managed by a host with fewer properties may obtain greater benefits (Xie et al., 2019). It has been found that property quantity and seller popularity as two trust-building signals can affect each other and influence customer purchases (Xie et al., 2019). Overall, customers tend to choose a host characterized by fewer listings and more historical sales because such hosts are trusted when purchasing a product.

As a signal with a persuasive effect, review valence not only directly affects customer purchases but also interacts with trust-building signals in influencing customer purchases (Ye et al., 2019; Yi et al., 2021). Specifically, review valence can work with the signals from product descriptions and provider profiles to affect sales performance (Wang et al., 2016). In online platforms, marketer- and customer-generated contents constitute the platform signals, which can collectively promote customer purchases under information asymmetry (Liang et al., 2020). Accordingly, we argue that positive review valence for service providers characterized by few listings and more historical sales are trusted by customers because the phenomena competitive hosts gain positive peer evaluation are authentic. Thus, we hypothesize that

Hypothesis 5a: The positive effect of review sentiment given to sellers who own fewer listings to achieve more historical sales on sales performance is stronger than that to those who own multiple listings or achieve fewer historical sales.

Hypothesis 5b: The positive effect of average rating given to sellers who own fewer listings to achieve more historical sales on sales performance is stronger than that to those who own multiple listings or achieve fewer historical sales.

Figure 3.1 presents our research framework to examine the effect of review valence (i.e., review sentiment and average rating) on sales performance and the moderating role of property quantity and seller popularity in this relationship.

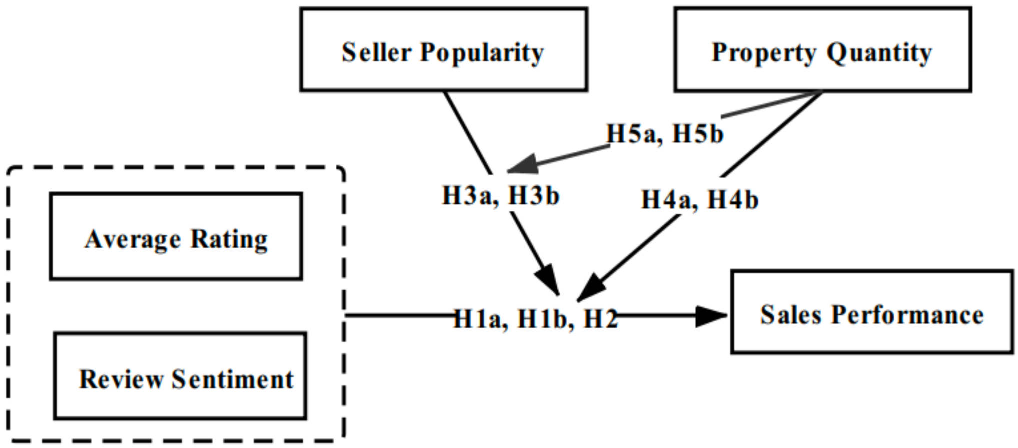


Figure 3.1 Research framework for Study 1

3.4 Data and Methodology

3.4.1 Dataset and sample

Xiaozhu.com is one of the leading online short-term rentals for accommodation in China, offering over 800,000 listings in more than 700 cities with about 50 million active users[6]. Given the explosive popularity in the sharing economy, scholars have expressed a preference for Xiaozhu as a representative P2P accommodation operator in the country (Wu et al., 2017; Xie et al., 2019). In addition, this platform aims to create a brand that appeals for humane services and social interactions. Moreover, the number of reviews is growing exponentially, shown in Figure 3.2, which can help analyze the signals from customer-generated content. Using python to crawl the HTML, we collected 102,335 online reviews, 29,994 listings, and 12,275 hosts information from 2012 to 2016, covering 17 cities. In our research context, we aim to examine the time-

[6] https://www.xiaozhu.com (accessed on July 28, 2020)

lagged effect of review valence on sales performance. Accordingly, the selected sample before 2016 should have more than three customer reviews, and we can obtain the metric of average rating. In addition, the selected sample should still operate in the market in 2016 and have customer reviews to represent sales performance. Moreover, we match all the information and delete the sample without detailed host profiles (e.g., historical sales volume). Finally, 3,260 listings remain for our subsequent observation.

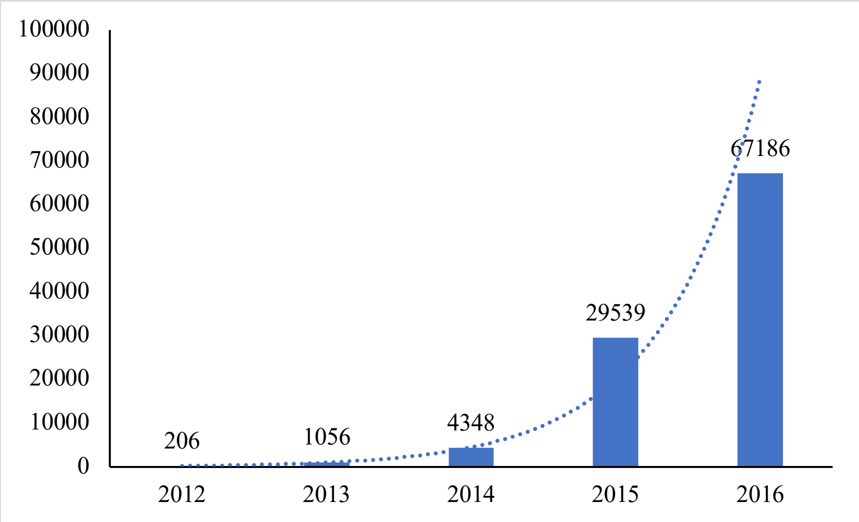


Figure 3.2 Number of reviews from 2012 to 2016 in Xiaozhu

Table 3.1 Data from 2012 to 2016 in seventeen cities

City	Number of listings	Number of providers	Number of reviews
Beijing	986	537	20881
Shanghai	417	267	6597
Chengdu	280	176	4742
Qingdao	238	138	2889
Xian	191	117	3370
Chongqing	157	93	2513
Guangzhou	135	88	2571
Xiamen	137	74	1643
Hangzhou	88	60	1761
Shenzhen	94	59	1365
Qinhuangdao	125	76	1521
Sanya	144	48	1523
Wuhan	91	52	1152
Suzhou	71	41	962
Dalian	87	38	814
Dali	10	6	114
Lijiang	9	7	140

3.4.2 Research methodology

To investigate the effects of review valence on sales performance, we employ text mining to capture the sentiment implied in online customer reviews. To conduct text analysis, we adopt machine learning classification algorithms to train the model and select Naïve Bayes (NB) as the optimal classifier to extract the sentiment for each review. Next, we use econometric models to empirically gauge the influence of review valence (i.e., review sentiment and average rating) on sales performance and test the moderating role of seller popularity and property quantity on this relationship.

(1) Text analysis

Text analysis is a research approach that depicts "what" is said and "how" it is expressed using qualitative and quantitative methods (Chang et al., 2020). It aims to understand the core meaning the text senders convey and predict the impact of the text on receivers' behaviors (Berger et al., 2020). To calculate review sentiment, we judge the sentiment by figuring out whether a review contains negative complaints. As a semi-supervised method, machine learning algorithms can help extract the sentiment from a large amount of unstructured data (Berger et al., 2020). It is an effective way to help humans deal with laborious tasks quickly by implementing entity extraction. In addition, the sentiment of online customer reviews is commonly divided into two categories, positive and negative (Chang et al., 2020). Thus, we adopt a binary classifier system to complete the machine learning process.

The processes of sentiment extraction are divided into four steps, shown in Figure 3.3. The first step is to preprocess data. According to prior studies (Berger et al., 2020; Chang et al., 2020), we break text into sentences (tokenization), remove nonmeaningful text (cleaning), and segment sentences into words or phrases using stopwords[7] and user-defined corpus extracted from our research context after acquiring online reviews. Second, text presentation is a process transited from human language to machine-recognized language, which can be regarded as model construction. We filter the bags of segmented words to extract features and then construct the bag-of-words-based model (CBOW) using a word-document matrix weighted by the word frequency. Third, we randomly select 5348 online reviews and label them "negative/positive" based on the sentiment of a review as training and testing samples (the ratio is 3:1). We adopt several machine learning classification algorithms to predict the sentiment and compare the prediction performance. According to the precision rate, recall rate, and F1 score, as well as micro-average, macro-average, and weighted-average metrics (Chang et al., 2020), we find that the Naïve Byes (NB) algorithm performs better than other classifiers in sentiment extraction, illustrated in Figure 3.4. Furthermore, we also calculate the accuracy and the Area Under the Curve (AUC) to address the imbalance in measuring

7 https://github.com/goto456/stopwords (accessed on 14 August 2020)

classification performance based on the Receiver Operating Characteristic Curve (ROC Curve), which illustrates the diagnostic ability of a binary classifier system[8]. All the results support that the classification performance of the NB algorithm is well-accepted (Accuracy=0.9, AUC=0.8107). Finally, we use the NB algorithm to predict the sentiment of e-WOM for 102335 online reviews and adopt the average sentiment scores at the listing level to measure review sentiment for each listing.

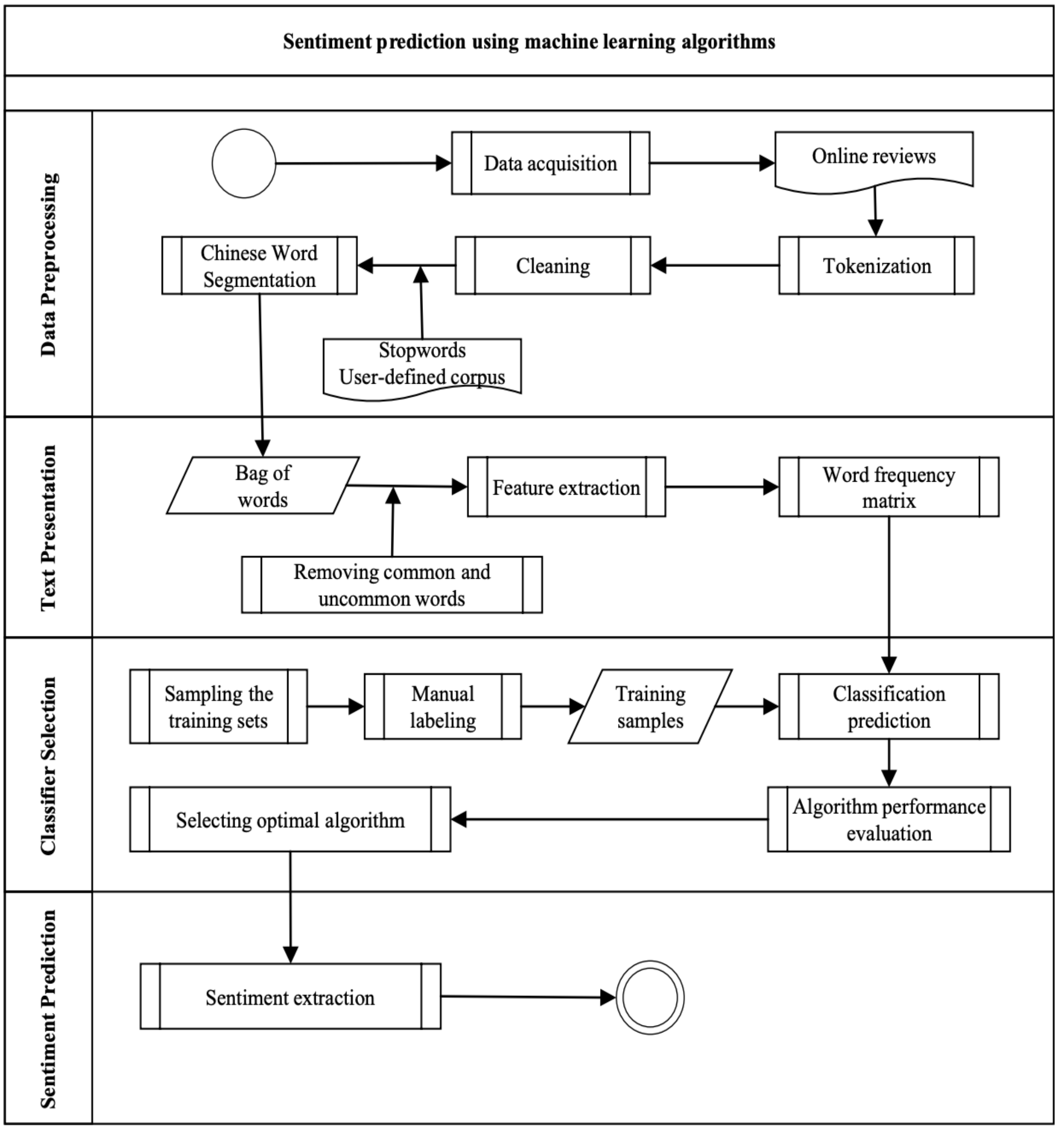


Figure 3.3 Sentiment extraction based on machine learning algorithms

8 https://en.wikipedia.org/wiki/Receiver_operating_characteristic (accessed on 14 August 2020)

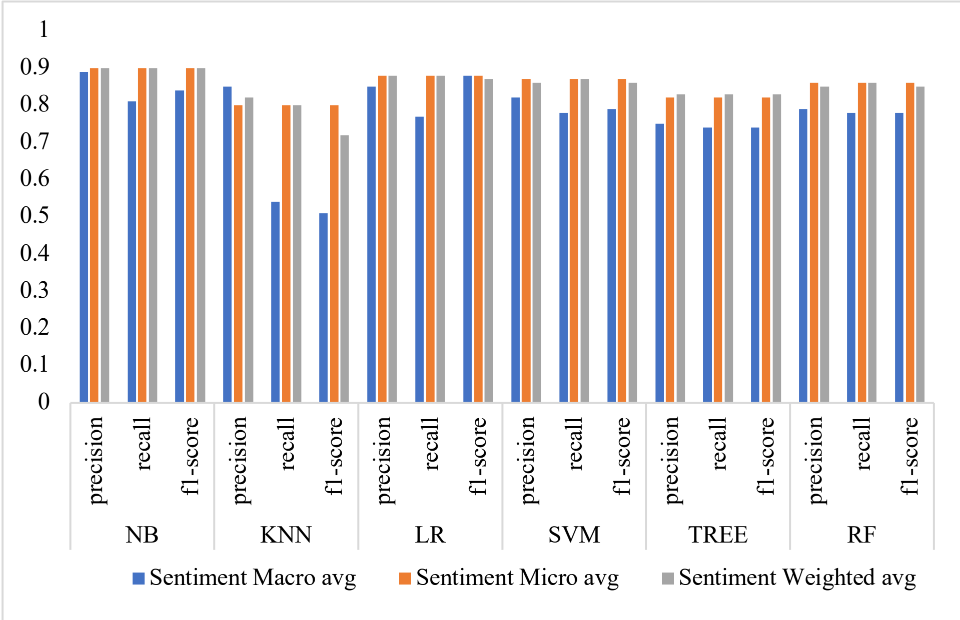


Figure 3.4 Classification performance of sentiment prediction

(2) Empirical analysis

To test our hypotheses, we adopt econometric models to empirically investigate the effect of review valence on sales performance and the moderating role of seller popularity and property quantity in this relationship. Considering the difficulty of obtaining the actual sales volume for each listing, we use the number of reviews as the proxy of sales performance (Liang et al., 2020). In terms of review valence, we use review sentiment and average rating as its indexes (Chen et al., 2017). The former is measured by the average sentiment score extracted from online customer reviews, and the latter is obtained from rating scores displayed on online platforms. Following previous literature (Chen, 2008; Duan et al., 2009), we adopt the historical sales volume that a seller achieves as the index of seller popularity. In addition, we measure property quantity by calculating the number of listings that a host owns (Xie et al., 2019).

To ensure empirical rigor, we include a robust set of control variables for testing our research model. First, considering customer preferences for various listing attributes (Xie et al., 2019; Yi et al., 2021), we add *Price, Area, Bedroom,* and *Bathroom* into the model. Second, due to the importance of host characteristics (Wu et al., 2017), we include *Response Rate, Confirmation Time,* and *Acceptation Rate* as control variables. Third, evidence has shown that previous performance will affect subsequent outcomes (Ke & Acm, 2017), indicating the rich-get-richer mechanism. We consider the *Number of Reviews* in the previous period as a control variable. Finally, we introduce the dummy variables of *City* and *Genre* controlled in our proposed model.

Table 3.2 presents the descriptive statistics, and a correlation matrix is shown in Table 3.3. The correlation matrix indicates no severe collinearity problem because the absolute value of the correlation coefficient ranges from 0.003 to 0.784. Moreover, we have conducted the variance inflation factor (VIF) test in subsequent regression analyses. All the results indicate no serious multicollinearity issue because the value of VIF is far less than 10.

Table 3.2 Variable description and statistics in Study 1

Variable	Definition	Mean	S. D.	Min	Max
Sales Performance	Number of reviews a listing receives in 2016	7.690	9.515	1	73
Number of Reviews	Number of reviews a listing receives before 2016	7.794	9.445	1	96
Review Sentiment	Average sentiment score of all reviews that a listing receives before 2016	0.791	0.401	-1	1
Average Rating	Average rating that customers give before 2016 on a scale of 1-5, with values of 1=terrible, 2=poor, 3=average, 4=good, and 5=excellent	4.158	0.771	1	5
Property Quantity	Total number of properties that a host owns	6.746	7.767	1	77
Seller Popularity	The number of reservations that a host sells	233.054	318.749	3	2002
Price	Average price of a listing	294.064	286.243	40	4800
Area	Listing area in square meters (m^2)	58.637	59.344	2	1600
Bedroom	Number of bedrooms	2.121	1.242	1	12
Bathroom	Number of bathrooms	1.285	0.725	0	12
Gender	A dummy variable with the value of 0=female and 1=male	0.333	0.471	0	1
Response Rate	Number of host responses versus the number of renters inquires (i.e., the rate that a host responds to questions asked by renters)	0.941	0.079	0.24	1
Confirmation Time	The minutes it takes for a host to confirm the renter reservation	5.156	7.659	0	340
Acceptance Rate	The acceptance rate of renter reservations	0.861	0.121	0.1429	1
City	A category variable about 17 cities in China	6.981	5.564	1	17
Genre	Room types on a scale of 1-3, 1=entire apartment, 2=private room, 3=shared room	1.332	0.513	1	3

Table 3.3 Correlations in Study 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Sales Performance	1.000															
(2) Number of Reviews	0.466	1.000														
(3) Review Sentiment	0.091	0.060	1.000													
(4) Average Rating	0.235	0.226	0.171	1.000												
(5) Property Quantity	-0.139	-0.103	-0.142	-0.127	1.000											
(6) Seller Popularity	0.172	0.217	-0.104	0.037	0.628	1.000										
(7) Price	-0.053	-0.058	-0.010	-0.036	-0.040	-0.099	1.000									
(8) Area	-0.050	-0.060	-0.021	-0.038	-0.028	-0.131	0.784	1.000								
(9) Bedroom	-0.039	-0.039	0.094	0.048	0.024	0.066	0.279	0.260	1.000							
(10) Bathroom	-0.035	-0.042	0.040	0.022	-0.036	-0.025	0.365	0.314	0.647	1.000						
(11) Gender	-0.039	-0.023	-0.014	-0.059	-0.006	-0.041	0.010	0.029	-0.015	-0.023	1.000					
(12) Response Rate	0.194	0.133	0.094	0.153	0.059	0.089	0.031	0.035	0.048	0.021	-0.070	1.000				
(13) Confirmation Time	-0.059	-0.043	-0.013	-0.084	-0.096	-0.071	0.033	0.028	0.023	0.022	-0.011	-0.175	1.000			
(14) Acceptance Rate	0.253	0.201	0.032	0.185	0.069	0.149	-0.065	0.010	0.012	-0.005	-0.065	0.427	-0.218	1.000		
(15) City	-0.132	-0.119	-0.054	-0.084	-0.029	-0.166	-0.038	0.054	-0.014	0.048	0.025	0.037	-0.008	0.052	1.000	
(16) Genre	0.017	0.044	0.137	0.109	0.024	0.167	-0.321	-0.419	0.434	0.177	-0.055	-0.003	0.037	-0.070	-0.147	1.000

Note: The absolute value is not less than 0.028, $P<0.1$; The absolute value is not less than 0.035, $p<0.05$; The absolute value is not less than 0.048, $P<0.01$.

To investigate the effect of review valence as well as the interactions with seller popularity and property quantity on sales performance, we model *Sales Performance* output as a function of the input of *Review Sentiment*, *Average Rating*, moderators, the interaction items, and control variables. To make variables comparable, we standardize all the variables. The estimation equation for each listing i is as follows.

$$
\begin{aligned}
&Sales\ Performance_{it} \\
&\quad = \beta_0 + \beta_1 Number\ of\ Reviews_{it-1} + \beta_2\ Review\ Sentiment_{it-1} \\
&\quad + \beta_3 Average\ Rating_{it-1} + \beta_4 Property\ Quantity_{it-1} \\
&\quad + \beta_5 Seller\ Popularity\ _{it-1} \\
&\quad + \beta_6 Review\ Sentiment_{it-1} \times Property\ Quantity_{it-1} \\
&\quad + \beta_7 Average\ Rating_{it-1}\ \times Property\ Quantity_{it-1} \\
&\quad + \beta_8 Review\ Sentiment_{it-1}\ \times Seller\ Popularity_{it-1} \\
&\quad + \beta_9 Average\ Rating_{it-1} \times Seller\ Popularity_{it-1} \\
&\quad + \beta_{10} Property\ Quantity_{it-1}\ \times Seller\ Popularity_{it-1} \\
&\quad + \beta_{11} Review\ Sentiment_{it-1} \times Property\ Quantity_{it-1} \\
&\quad \times Seller\ Popularity_{it-1} \\
&\quad + \beta_{12} Average\ Rating_{it-1} \times Property\ Quantity_{it-1} \\
&\quad \times Seller\ Popularity_{it-1} + \beta_{13} HOST_{it-1} + \beta_{14} LISTING_{it-1} + \beta_{15} City_i \\
&\quad + \beta_{16} Genre_i + \varepsilon_{it}{}'
\end{aligned}
$$

Where HOST is a vector of host attributes controls that include *Gender*, *Response Rate, Confirmation Time, and Acceptation Rate*. LISTING is a vector of listing characteristics that consist of *Price, Area, Bedroom,* and *Bathroom.* In addition, we introduce a dummy variable *City* and *Genre* as the proxy of the city of samples and the index of room type, respectively. ε' is the random error.

3.5 Results and Analyses

3.5.1 Main effects

Table 3.4 presents the standardized estimation of review valence and its interactions with seller popularity and property quantity on sales performance. Model 1 includes review valence (i.e., review sentiment and average rating), moderators (i.e., property quantity and seller popularity), and control variables, whereas Model 2 adds the two-way interactions of review valence with property quantity or seller popularity, respectively. Model 3 adds the three-way interactions of review valence with property quantity and seller popularity.

Models 1, 2, and 3 from Table 3.4 have consistently suggested that review sentiment (β=0.079, p<0.001) and average rating (β=0.117, p<0.001) are positively associated with sales performance. Thus, Hypotheses 1a and 1b are supported. It is consistent with previous literature arguing that review valence can significantly affect customer purchase because it can illustrate customer opinions and satisfaction (Jia & Liu, 2018).

To compare the effects of review sentiment and average rating on sales performance, we refer to previous research (Cohen et al., 2013) and conduct the difference test between their standardized coefficients (t= $\frac{\beta_3-\beta_2}{SE_{\beta_3-\beta_2}}$ = (0.117-0.079)/ $\sqrt{\frac{1-0.3379}{3221} \times [1.3790 + 1.7553 - 2 \times (-0.2668)]} = 1.38 < 1.96$, df=3221, n.s.). The statistic demonstrates no substantial difference between review sentiment and average rating in affecting sales performance. Thus, contrary to our expectation, Hypothesis 2 is not supported. One possible explanation is that review sentiment can better illustrate online rating (Zhu et al., 2020).

Table 3.4 Standardized estimation of sales performance in Study 1

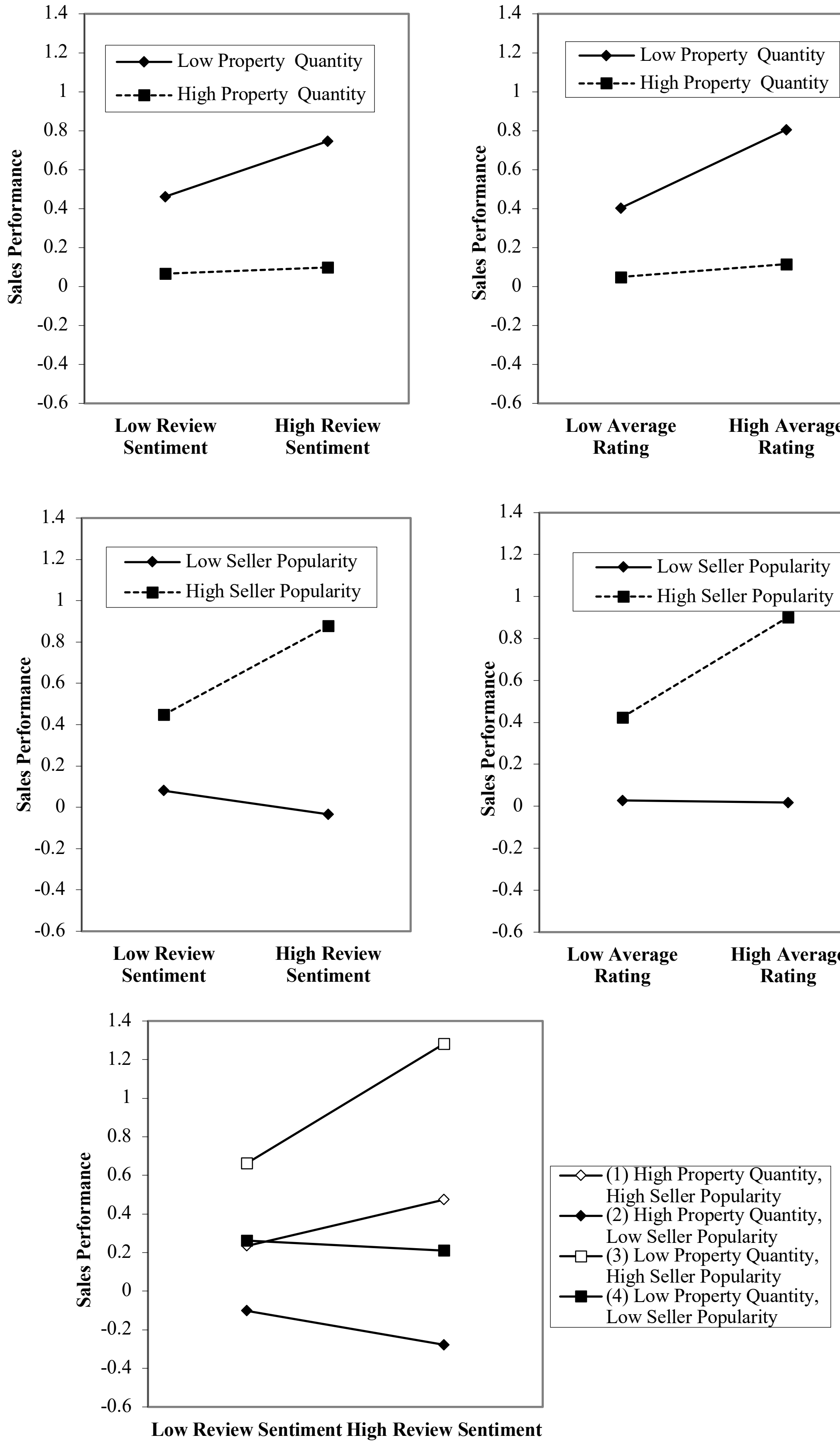
Sales performance	Model 1	Model 2	Model 3
Number of Reviews	0.324***(0.016)	0.304***(0.016)	0.289***(0.017)
Review Sentiment	0.038*(0.015)	0.055***(0.016)	0.079***(0.017)
Average Rating	0.076***(0.016)	0.105***(0.017)	0.117***(0.019)
Property Quantity	-0.246***(0.021)	-0.289***(0.022)	-0.261***(0.023)
Seller Popularity	0.220***(0.021)	0.257***(0.022)	0.320***(0.027)
Review Sentiment × Property Quantity		-0.073***(0.017)	-0.063**(0.018)
Average Rating × Property Quantity		-0.101***(0.019)	-0.084***(0.019)
Review Sentiment × Seller Popularity		0.080***(0.020)	0.136***(0.026)
Average Rating × Seller Popularity		0.119***(0.024)	0.122***(0.030)
Property Quantity× Seller Popularity			-0.048***(0.009)
Review Sentiment × Property Quantity × Seller Popularity			-0.032***(0.009)
Average Rating × Property Quantity × Seller Popularity			-0.030**(0.011)
Price	-0.051+(0.027)	-0.047+(0.027)	-0.045+(0.027)
Area	0.007(0.028)	-0.001(0.028)	0.007(0.028)
Bedroom	-0.0004(0.025)	-0.00007(0.024)	-0.006(0.024)
Bathroom	0.008(0.020)	0.012(0.020)	0.010(0.020)
Gender	-0.003(0.015)	-0.004(0.015)	0.005(0.015)
Response Rate	0.077***(0.017)	0.072***(0.016)	0.070***(0.016)
Confirmation Time	0.0002(0.015)	-0.002(0.015)	-0.0002(0.015)
Acceptance Rate	0.124***(0.017)	0.124***(0.017)	0.119***(0.017)
City Dummies	Included	Included	Included
Genre Dummies	Included	Included	Included
Intercept	0.299(0.278)	0.314(0.275)	0.343(0.274)
F	48.08***	45.35***	43.25***
R-squared	0.3159	0.3299	0.3379
Mean VIF	1.62	1.70	2.09

Note: N=3260, Standardized *b*'s are reported with robust standard errors in parentheses.
+p<0.1, *p < .05, **p < .01, ***p < .001

3.5.2 Interaction effects

Models 2 and 3 from Table 3.4 illustrate the moderating role of seller popularity and property quantity between review valence and sales performance. Evidence has shown that seller popularity strengthens the positive effects of review sentiment (β=0.136, $P<0.001$) and average rating (β=0.122, $P<0.001$) on sales performance. The positive impact of review valence given to hosts achieving more historical sales on sales performance is stronger than that given to hosts achieving few historical sales. These findings support Hypotheses 3a and 3b. It confirms the herd behavior in online shopping, indicating that customer reviews and choices as two signals of popularity can affect each other and collectively determine customer purchases (Chen, 2008; Cheung et al., 2014). In addition, the positive effects of review sentiment (β=-0.063, $P<0.01$) and average rating (β=-0.084, $P<0.001$) on sales performance are weakened by property quantity that a host owns. In other words, the positive effect of review valence given to regular hosts on sales performance is stronger than that given to multi-listings hosts. Thus, hypotheses 4a and 4b are supported. These findings are consistent with previous studies supporting that property quantity may negatively affect sales performance (Xie et al., 2019) and weaken the relationship between trust-building signals and customer purchases (Liang et al., 2020).

In terms of three-way interactions of review valence with seller popularity and property quantity, the positive effects of review sentiment (β=-0.032, $P<0.001$) and average rating (β=-0.030, $P<0.01$) given to hosts who own fewer listings to achieve more historical sales on sales performance are more prominent than that given to hosts who own multiple listings or achieve fewer historical sales. Thus, hypotheses 5a and 5b are supported. Our findings are in line with prior literature suggesting the negative moderating role of listings volume and the positive moderating role of historical sales volume in the relationship between trust-building signals and sales performance (Cheung et al., 2014; Xie et al., 2019). Figure 3.5 illustrates the two-way and three-way interactions of review valence (i.e., review sentiment and average rating) with property quantity and seller popularity on sales performance, respectively and collectively.


Sales Performance
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Low Property Quantity
High Property Quantity
Low Review Sentiment
High Review Sentiment
Low Average Rating
High Average Rating
Low Seller Popularity
High Seller Popularity
(1) High Property Quantity, High Seller Popularity
(2) High Property Quantity, Low Seller Popularity
(3) Low Property Quantity, High Seller Popularity
(4) Low Property Quantity, Low Seller Popularity

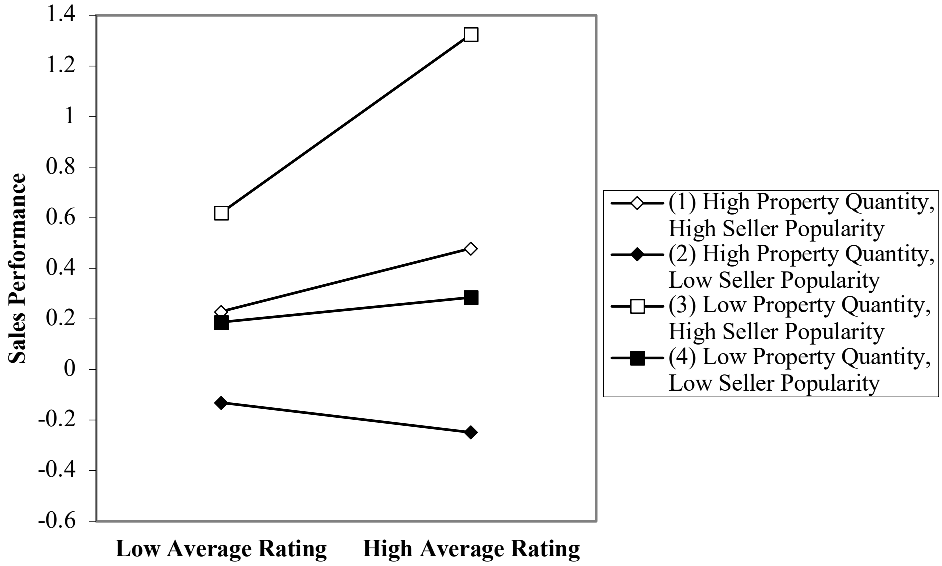


Figure 3.5 Moderating role of property quantity and seller popularity

3.5.3 Robustness check

We have taken multiple precautions to ensure robustness. For example, we adopt two alternative measures of review valence, such as average rating and review sentiment, to support our proposed model. In addition, we sample seventeen cities and three types of products to test our research model and use city and genre fixed effect controlled in this model. Moreover, we also conduct various tests to check the robustness of our findings. Table 3.5 reports the results of robustness tests. First, we add a dummy variable of whether a host discloses the personal homepage in online platforms into our original model because information disclosure may affect customer purchases (Xie et al., 2019), as shown in Model 1. Second, we change *Price* into a categorical variable with the value of low (Price<168), medium (168<=Price<299, and high (299<=Price) based on the distribution of *Price*, as presented in Model 2. Third, prior literature has argued that customers value various attributes concerning the product or service providers staying at different sharing-level listings (Xu, 2020). We replace the dummy variable of *Genre* with the categorical variable of *Sharing level* because various listings mean the space sharing at different levels, as shown in Model 3. Finally, considering the measurement errors from text mining (Yang et al., 2018), we introduce 10% errors through human intervention by changing the prediction results to the opposite ones, as illustrated in Model 4. In summary, all the results are consistent with the findings in our proposed model and help rule out some alternative explanations of our research.

Considering that review sentiment is extracted from online customer reviews, there may exist measurement bias for review sentiment. In addition, seller popularity and property quantity depict the signal of service quality, which may influence online customer reviews. Overall, review sentiment is endogenous because it may be affected by the moderators. The coefficients of interaction terms between moderators and review sentiment using OLS would be inflated (Zhang et al., 2018). Following previous literature (Hamilton & Nickerson, 2003), we use two-stage least squares (2SLS) regressions to overcome the endogeneity problems. In the first stage, we regress review sentiment on seller popularity, property quantity, and other control variables to obtain the residual. The result of the stage-one estimate indicates that seller popularity (β=-0.112, P<0.001) and property quantity (β=-0.065, P<0.01) are significantly associated with review sentiment, supporting the use of 2SLS to correct for potential endogeneity issues. In the second stage, we use the residual as the index of review sentiment to test our proposed model, shown in Model 5. The results are consistent with our findings in Table 3.4, indicating the robustness of our research framework. To avoid reversal causality, we adopt one year lag between review valence and sales performance in our original model. Moreover, we also include previous review volume and many essential variables from the previous literature controlled in this model. Overall, it shows that after addressing the endogenous problem, the main effects are supported. In addition, we provide evidence to support the relationship between review valence and sales performance from the perspective of causality.

Table 3.5 Robustness check of Study 1

Sales performance	Model 1	Model 2	Model 3	Model4	Model 5
Number of Reviews	0.283***(0.017)	0.290***(0.017)	0.289***(0.017)	0.295***(0.017)	0.297***(0.016)
Review Sentiment	0.077***(0.017)	0.079***(0.017)	0.077***(0.017)	0.067***(0.017)	0.064***(0.016)
Average Rating	0.116***(0.019)	0.117***(0.019)	0.116***(0.019)	0.123***(0.019)	0.133***(0.019)
Property Quantity	-0.264***(0.023)	-0.261***(0.023)	-0.259***(0.023)	-0.252***(0.023)	-0.249***(0.023)
Seller Popularity	0.311***(0.027)	0.323***(0.027)	0.320***(0.027)	0.305***(0.027)	0.285***(0.026)
Review Sentiment × Property Quantity	-0.064**(0.018)	-0.064***(0.018)	-0.063***(0.018)	-0.047*(0.019)	-0.044*(0.018)
Average Rating × Property Quantity	-0.085***(0.019)	-0.084***(0.019)	-0.082***(0.019)	-0.090***(0.019)	-0.096***(0.019)
Review Sentiment × Seller Popularity	0.133***(0.026)	0.135***(0.026)	0.136***(0.026)	0.099***(0.028)	0.093***(0.026)
Average Rating × Seller Popularity	0.121***(0.030)	0.121***(0.030)	0.120***(0.030)	0.135***(0.030)	0.154***(0.029)
Property Quantity × Seller Popularity	-0.044***(0.009)	-0.048***(0.009)	-0.048***(0.009)	-0.046***(0.009)	-0.039***(0.009)

Review Sentiment × Property Quantity × Seller Popularity	-0.030**(0.009)	-0.032***(0.009)	-0.032***(0.009)	-0.026**(0.009)	-0.022*(0.009)
Average Rating × Property Quantity × Seller Popularity	-0.029**(0.011)	-0.030**(0.011)	-0.030**(0.011)	-0.034**(0.011)	-0.038**(0.011)
Price	-0.047+(0.027)		-0.048+(0.027)	-0.047+(0.027)	-0.047+(0.027)
Area	0.007(0.028)	-0.026(0.021)	0.030(0.027)	0.011(0.028)	0.010(0.028)
Bedroom	-0.004(0.024)	-0.010(0.025)	-0.028(0.023)	-0.006(0.024)	-0.003(0.024)
Bathroom	0.009(0.020)	0.004(0.020)	0.011(0.020)	0.008(0.020)	0.008(0.020)
Gender	0.006(0.015)	0.005(0.015)	0.005(0.015)	0.003(0.015)	0.004(0.015)
Response Rate	0.069***(0.016)	0.069***(0.016)	0.070***(0.016)	0.073***(0.016)	0.074***(0.016)
Confirmation Time	-0.0003(0.015)	0.0002(0.015)	-0.001(0.015)	0.0004(0.015)	-0.001(0.015)
Acceptance Rate	0.115***(0.017)	0.122***(0.017)	0.119***(0.017)	0.119***(0.017)	0.120***(0.017)
Homepage	0.047**(0.015)				
Price level		0.006(0.020)			
Sharing level			-0.059**(0.022)		
City Dummies	Included	Included	Included	Included	Included
Genre Dummies	Included	Included		Included	Included
Intercept	0.313(0.273)	0.355(0.273)	0.298(0.274)	0.354(0.274)	0.360(0.274)
F	42.5***	43.14***	44.13***	42.58***	42.57***
R-squared	0.3398	0.3373	0.3363	0.3344	0.3343
Mean VIF	2.05	1.96	2.15	2.06	2.04

Note: N=3260, Standardized *b*'s are reported with robust standard errors in parentheses.

+p<0.1, *p < .05, **p < .01, ***p < .001

3.6 Discussions

3.6.1 Findings

In e-commerce, where information asymmetry is obvious, review valence acts as a trust-building signal with the persuasive effect in swaying customer decisions (Duan et al., 2008). Drawing upon the signaling theory, we examine and compare the effect of review valence on sales performance. In addition, we adopt the contingency theory perspective to establish boundary conditions on the effectiveness of this relationship. Using a combination of text mining and econometric analysis, we provide evidence supporting the positive effect of review valence (i.e., review sentiment and average rating) on sales performance. It further confirms that review valence is crucial for customers' judgment over the product, determining customer behavior (Floyd et al., 2014; You et al., 2015). Notably, we find that review sentiment does not differ from average rating in influencing sales performance. It demonstrates that review sentiment signals a persuasive effect in the same way as average rating (Zhu et al., 2020).

In terms of interaction effects, the positive effect of review valence on sales performance is weakened by property quantity. Multi-listing hosts face insufficient resource allocation and capability, as they should balance the trade-off between service quality and listing quantity (Ellway, 2014). On the other hand, we provide evidence to support that seller popularity strengthens the positive effects of review valence on sales performance. It also indicates that the popularity-related signals can convey product and service quality, which is consistent with previous literature (Cheung et al., 2014). More importantly, we find that the three-way interactions of review valence (i.e., review sentiment and average rating) with property quantity and seller popularity is significant in determining sales performance. Specifically, the positive impact of review valence given to hosts who own fewer listings to achieve more historical sales on sales performance is more prominent than that given to hosts who own multiple listings or achieve fewer historical sales. These findings mirror that service providers characterized with personalization and popularity convey more trustworthiness (Chen, 2008; Ju et al., 2019). Overall, online customer reviews can work together with provider profiles to significantly affect customer purchases.

3.6.2 Theoretical implications

As a community-based service, the sharing economy emphasizes the importance of trust-building signals between sellers and buyers (Ert et al., 2016; Wu et al., 2017). Our research contributes knowledge to the sharing economy in three aspects. First, this study utilizes the signaling theory and observational learning theory to explain customer behavior in a platform-mediated economy. As a reliable information source, online customer reviews convey trustworthy signals to potential customers (Mariani & Borghi, 2020). Our research has supported that review valence can play a persuasive role in influencing customer behavior (You et al., 2015). In addition, we have further confirmed that host attributes can carry helpful information for customers' choices (Wu et al., 2017). In particular, herd behavior is evident in online shopping because popularity can positively affect customer purchases (Chen, 2008). Overall, online reviews and host profiles can contribute to sales performance due to their trust-building signals in the presence of information asymmetry. Our findings extend the signaling theory and verify the observational learning theory by illustrating the interactions of the signals from online reviews and host profiles in affecting sales performance in the sharing economy.

Second, this study highlights the importance of review valence by demonstrating their positive effects on sales performance in the platform-mediated economy. In this study, we consider not only review sentiment but also average rating to measure review valence, indicating that unstructured data is as crucial as structured one when conveying customer opinions. It confirms that review valence plays an essential role in swaying customer purchase decisions because it can better reflect customer experience (Floyd et al., 2014; You et al., 2015). Moreover, it answers the paradox about the effect of review valence on sales performance with empirical evidence (Duan et al., 2008; Jia & Liu, 2018). In addition, we add new knowledge by uncovering the equal importance of review sentiment and average rating, as two manifestations of review valence, in determining sales performance. This finding supports that textual sentiment is an effective way to predict online review rating in online platforms (Li et al., 2019; Zhu et al., 2020).

Third, we examine the moderating effects of property quantity and seller popularity on the relationship between review valence and sales performance. Specifically, it highlights the interactions of online reviews with host attributes in determining sales performance. In the absence of trust, customers tend to trust collective opinions and choices to reduce risks from uncertain product or service quality. This study demonstrates the complementary effect of seller popularity on the link between review valence and sales performance, indicating that customers are likely to follow others' choices and trust service providers characterized with popularity (Cheung et al., 2014). Moreover, we have provided evidence supporting the negative effect of property quantity on sales performance. Our research contrasts with that of the traditional transaction, where customers value the professionalism and experience of service providers (Wu et al., 2017). In the sharing economy, the allocation of capacities and resources for multi-listing hosts to each listing is limited (Xie & Mao, 2017). It indicates that compared to multi-listing hosts, service providers with fewer listings are more favored by the market. The latter has more resources (i.e., time and energy) to serve customers and offer quality service (Xie & Mao, 2017; Liang et al., 2020). Our finding addresses widely disputed issues on property quantity, which can offer novel ideas on customer behaviors in the sharing economy. More importantly, we have supported the three-way interactions of review valence with property quantity and seller popularity in affecting sales performance. It further provides clues supporting the interactions of multiple trust-building signals in influencing sales performance in the sharing economy (Xie et al., 2019). In particular, the positive effect of review valence given to hosts who own fewer listings to achieve more historical sales on sales performance is more prominent than that given to hosts who own multiple listings or achieve fewer historical sales. These findings enhance our knowledge about the boundary conditions of promoting the impact of review valence on sales performance in service operations management.

3.6.3 Practical implications

Our findings give some suggestions for service providers to optimize their business model and improve their service operations management in the sharing economy. Regarding the importance of review valence, it is highly desirable for the platform to design the topic about average rating and review sentiment to appear in an eye-catching position. To attract and retain customers, service providers are advised to satisfy customers through offering quality services and hospitality practices, motivating customers to voluntarily post positive reviews and give a high score rating on service and product quality. As for service providers who have a low reputation manifesting in review sentiment and average rating, they need to balance their capacity to manage products and offer high-quality service. Specifically, it is suggested that they should reduce the listing volume managed by them, which helps them improve their abilities to offer personalized service. In addition, service providers are suggested to launch various promotional strategies, such as giving discounts or gifts, to increase sales and achieve their popularity. Overall, service providers should strive to obtain positive e-WOM by offering consumer-oriented service to strengthen the customer-brand relationship. Moreover, it is desirable for them to balance product quantity with quality in their efforts to enhance sales performance.

To sustain the sharing economy, we have provided some ideas on regulating and advancing its development. Considering that review valence helps offer clues for potential customers, the market operators should encourage customers to actively evaluate service and product quality by posting their opinions. In contrast, the market operators should put a ban to review manipulation (e.g., fake reviews) and supervise the review management. In the sharing economy, service quality is equally important as product quality, and even the former may take precedence over the latter. The market operators should invest more energies and resources in improving the overall quality of service providers to bolster the sustainable development of the sharing economy. For example, they are suggested to regularly provide opportunities for employee training and then conduct quality assessments of working performance among employees. Furthermore, the market supervisors should appropriately raise the entry barrier in selecting service providers according to their effort put into customer services. More importantly, the market operators can implement an incentive to offer ordinary service providers proper guidance and award service providers characterized with popularity, which can help energize the industry development. To create more profits, the market operators should give suggestions for multi-listing hosts to enhance their service operations management by better cooperation with the peers to allocate their resources and capabilities to each product. Doing so can create a win-win situation for service providers and their partners.

3.6.4 Limitations and future research

This study is subject to several limitations for interpretation of the study results. First, due to the phenomenon that service providers enter or exit the market frequently, it is challenging to capture multi-period data about online customer reviews to conduct panel analysis. Future research can extend our research model to other fields, e.g., shipping and port services, to capture a large number of samples and adopt panel analysis to verify the model. Specifically, scholars can extract review sentiment implied in customer feedback and link it to customer satisfaction in the liner shipping industry (Hirata, 2020). In addition, scholars can also identify the negative topics and summarize the risks in the port services (Lai et al., 2020). Second, we focus on the products with customer reviews, and sample selection bias can be present. Future research can combine lab experiments with surveys/interviews to obtain first-hand data and mitigate the sampling bias issue. Third, due to the difficulty of access to sales data, we consider review volume as the proxy of sales performance. It is helpful to cooperate with the business companies to obtain actual sales volume to triangulate our findings. Fourth, we focus on a Chinese accommodation-sharing platform. The results are less generalizable to other platforms (e.g., Airbnb and Tujia) and regions (e.g., the United States). It is worth extending this study to other platforms and examining individual differences among various accommodation-sharing platforms. A study on the cultural difference in choosing P2P accommodations by comparing customer behaviors from a specific platform in different regions or countries is also a promising research direction.

Chapter 4 Study 2: Self-presentation and Sales Performance: The Role of Customer-generated and Marketer-generated Reputations

4.1 Introduction

With the blooming of online channels, the platform-mediated economy has gradually risen to prominence as the bridge of information exchange between sellers and buyers (Belk, 2014). In the sharing economy, where customers may share a house or room with service providers after online booking, customers resort to multiple signals on online platforms and make their purchase decisions. To satisfy customers' needs for seeking helpful information, sellers are encouraged to disclose their product-related and personal information online (Ert et al., 2016). However, due to the absence of face-to-face communications before purchase, the information from sellers (senders) to buyers (receivers) is limited and asymmetric. Under this context, customers judge product quality from the visual cues displayed on online platforms to reduce uncertainties (Park & Nicolau, 2015). These cues are derived from customers, sellers, or third parties (e.g., the platform). Customer-generated content (online reviews) has been widely considered as a credible source for customer purchase decisions (Ayeh et al., 2013; Jia & Liu, 2018). Moreover, product description given by the platform or providers is also a decisive factor in influencing customer decision-making processes (Mauri et al., 2018). In particular, provider self-presentation has become a pivotal element that customers pay attention to when service providers participate in customer experience (García et al., 2019). In the accommodation-sharing economy, host-guest interaction is evident from online to offline channels. The interactive activities involve chatting, living in a shared room, and having dinner together. Undoubtedly, provider profiles are essential because customers put more value on host hospitality and social interactions (Tussyadiah & Zach, 2017; Belarmino et al., 2019).

As a piece of marketer-generated information, provider profiles have been attached central importance in the sharing economy (Wu et al., 2017; Liang et al., 2020). Especially for services in the exchange of money, customers pay more attention to service effort as an additional product (Zhao et al., 2015). It has been found that host attributes are the crucial factors determining customer trust (Wu et al., 2017). In addition, prior literature has suggested that self-presentation is positively associated with online social support (Lin et al., 2020), online impression management (Ellison et al., 2006), relationship maintenance and platform enjoyment (Krasnova et al., 2010), and information accessibility (Zhang et al., 2019). Most studies have starkly focused on social networking sites to investigate self-presentation (Ellison et al., 2006; Lin et al., 2020). Nevertheless, self-presentation as an advertising strategy to create profits in e-commerce is under-investigated. It needs an in-depth understanding to examine the outcomes of provider self-presentation, especially in the sharing economy context, where service providers participate in customer experience before, during, after customer encounter experience.

In the online environment, self-presentation conveys trust-building signals regarding service providers and has a positive association with customer choices (Ma et al., 2017). To advertise their products and obtain favorable responses, service providers start to build a seller-buyer relationship through strategic self-presentation (Tussyadiah, 2016). In the sharing economy, self-presentation is a form of host-guest interaction, which can stimulate customer resonance and behavioral intentions (Shang et al., 2017). Previous literature has found that provider self-presentation plays a crucial role in affecting customer trust and booking intention (Tussyadiah & Park, 2018). Specifically, service providers presenting them as travelers than workers can gain more popularity (Tussyadiah & Park, 2018). Scholars have echoed that social-oriented content can reduce social distance and build a close relationship between customers and service providers (García et al., 2019). Similarly, the openness embedded in self-presentation helps customers develop trust perception (Zhang et al., 2020). In addition, positive sentiment implied in host self-presentation can attract customers (Zhang et al., 2018). However, the studies investigating the effect of semantic features implied in provider self-presentation are scarce. Accordingly, this study aims to examine the impact of host self-presentation and its semantic feature (i.e., social-oriented vs. non-social-oriented) on sales performance in the accommodation-sharing economy.

Online reputation is a critical signal for customers to reduce perceived risks from information asymmetry (Belk, 2014). Following previous literature arguing that customer-generated and marketer/seller-generated content constitute the platform signals, there are two reputation mechanisms in online platforms: reputation evaluated by customers and ones given by platform systems (Liang et al., 2020). From this standpoint, we define the reputation generated from customers as customer-generated reputation. The reputation awarded by the platform is considered as marketer-generated reputation. For instance, customers can post a rating on product and service quality according to their consumption experience. Similarly, the platform offers a reputational badge to award service providers based on their performance. In the accommodation-sharing economy, star rating given by customers can be regarded as customer-generated reputation, and the badge of "superhost" granted by the platform can be considered as marketer-generated reputation (Biswas et al., 2020). Evidence has shown that reputation positively moderates the relationship between trust-building signals and sales performance in e-commerce (Wang et al., 2016). However, some scholars have argued that multiple signals may affect each other and even suppress (Xie et al., 2019). It motivates us to investigate how the effect of self-presentation on sales performance depends on customer-generated and marketer-generated reputations.

To pursue our research goals, we embrace the social distance theory and signaling theory to explain the interaction effects of self-presentation with reputation on sales performance. Social distance theory is derived from construal level theory arguing that individuals can form different levels of construal when experiencing events or objects, generating psychological distance that consists of temporal, social, and spatial distances (Trope & Liberman, 2010). It has been applied to explain that the interactions make customers feel psychologically close to service providers and affect their subsequent behavior in the online environment (Kim et al., 2008). Self-presentation is a way for service providers to help customers reduce social distance by building a virtual and digital self (Tussyadiah & Park, 2018). In particular, social-oriented self-presentation can arouse consumer resonance to perceive the intimacy from service providers (García et al., 2019). Moreover, as a signal of product brand, reputation is an essential factor swaying customer purchases because it conveys a trustworthy signal of service and product quality recognized by the public (Liang et al., 2020). Specifically, online reputation consists of marketer-generated reputation awarded by the platform (i.e., superhost) and customer-generated reputation given by customers (i.e., average rating).

In this study, we employ a combination of text mining and econometric analysis to figure out the relationship between self-presentation and sales performance as well as the moderating role of reputation in this relationship. In terms of text mining, we adopt latent Dirichlet allocation (LDA) to identify the topics of self-presentation and use the support vector machine (SVM) algorithm to extract semantic features of self-presentation, namely social orientation. LDA is an established approach to identify topics, while machine learning can effectively predict topics from unstructured text (Berger et al., 2020). These two methods are helpful to understand the signals from a large amount of text. For econometric analysis, we find that self-presentation, especially for social-oriented self-presentation, positively affects sales performance. Notably, the positive effects of self-presentation and social orientation on sales performance are substituted by marketer-generated reputation (i.e., superhost) and complemented by customer-generated reputation (i.e., review rating). Our findings provide new insights into the sharing economy literature by uncovering the prominence of self-presentation and the distinctive role of online reputation in moderating the relationship between self-presentation and sales performance. Practically, we offer helpful suggestions for market operators to formulate marketing strategies by constructing a virtual brand through self-presentation and gaining an online reputation and for policy makers on how to regulate and sustain the sharing economy.

4.2 Literature Review

4.2.1 Social distance

Construal level theory (CLT) illustrates that individuals perceive events or objects at different levels of concreteness or abstraction and generate psychological distance between themselves and their counterparts (Trope & Liberman, 2010). Psychological distance refers to "a subjective experience that something is close or far away from the self, here, and now". The distance can be derived from the temporal (in time), social (in social contexts), and spatial (in space) aspects. As one of the most crucial elements of psychological distance, social distance depicts how close or far individuals feel to other people, determining subsequent interactions (Stephan et al., 2011). It has been widely applied in the marketing, hospitality tourism, and information systems literature to explain customer behavior (Kim et al., 2008; García et al., 2019). Prior research has conceptualized social distance in the forms of interaction, interpersonal similarity, and dependence (Stephan et al., 2011). The reduction of social distance can help build customer trust because customers generate a close relationship with service providers (Tussyadiah & Park, 2018).

Due to the absence of face-to-face interaction in the digital age, the platform-mediated economy emphasizes the significance of constructing a solid psychological bond between sellers and buyers (García et al., 2019). In this context, social distance refers to the extent to which users interact with others whom they perceive to be psychologically close to themselves (Kwon, 2020). In the sharing economy, self-presentation is a way for service providers to interact with customers via online channels, making customers form a high construal level and feel close to service providers (Tussyadiah & Park, 2018). Evidence has shown that self-presentation in online peer-to-peer (P2P) platforms can reduce social distance and increase revenues (García et al., 2019). Specifically, the expressions of self-presentation can evoke social values and exerts a significant impact on the revenues (García et al., 2019). In particular, social-oriented self-presentation significantly influences customer purchases because it can make customers feel hospitable from service providers (Han et al., 2019). Based on the above, we consider self-presentation as the proxy of reducing social distance by constructing a digital self, which can affect sales performance

4.2.2 Provider self-presentation

Self-presentation is an advertising strategy for impression management in online platforms by creating a virtual profile that reflects themselves (Ellison et al., 2006). In online communities, self-presentation can disclose one's online identity and construct a digital self by conveying users' personality, experiences, and social attitudes (Jin et al., 2015). It can help build trust and strengthen an interpersonal relationship between senders and receivers of information when people disclose themselves to others (Gibbs et al., 2006). Previous studies have investigated the positive outcomes of self-presentation on impression management, customer trust, and customer purchases (Tussyadiah & Park, 2018; García et al., 2019). Overall, these studies have consistently supported that self-presentation conveys a positive signal of service quality because it can express service benevolence to help receivers make a proper judgment.

In the sharing economy, self-presentation is a crucial element of social interaction, which facilitates the matching between sellers and buyers in personality and hobbies, promoting customer purchase behavior (Moon et al., 2019). It is a strategic marketing attractiveness for service providers to obtain prospective customers because customers perceive a higher level of trustworthiness in hosts who disclose their information (Tussyadiah & Park, 2018). Through opening personal information, self-presentation is a way of expressing themselves and depicting their amateur, which helps reduce social distance (García et al., 2019). Through text mining from host descriptions, scholars have found two patterns of self-presentation in the accommodation-sharing economy: a well-traveled social enthusiast and an individual with a specific profession (Tussyadiah & Park, 2018). Specifically, a well-traveled host implied in host profiles can gain more customer trust and booking intention (Tussyadiah & Park, 2018). From the perspective of Aristotle's appeals, scholars have found that the use of social words significantly and positively affects the purchase (Han et al., 2019). In this study, we aim to investigate the effects of provider self-presentation and its semantic feature, such as social orientation, on sales performance in the accommodation-sharing economy.

4.2.3 Reputation

The reputation system reveals a significant trend for customers' decision support and successful commercial online applications in the platform-mediated economy (Josang et al., 2007). Online reputation is interpreted as a trustworthy signal of depicting service quality among potential customers, which can form a solid customer-brand relationship (Bertarelli, 2015). It can mirror general opinions and help customers trust a specific brand under the information asymmetry in online platforms (Liang et al., 2020). Moreover, online reputation not only motivates service providers to improve or professionalize their services but also gives customers heuristic clues for their purchase decisions (Liang et al., 2017). Evidence has shown that the reputation system helps customers reduce risks and make a proper judgment before purchasing (Chen et al., 2017). In addition, it has been found that online reputation can help practitioners achieve business success because it attracts customers to make a booking behavior (Liang et al., 2020). Besides, scholars have also found that guests are willing to pay more for "superhost" accommodations (Liang et al., 2017).

Online reputation can be represented by customer ratings (Bai et al., 2020) and the badge or status granted by the platform (Liang et al., 2017; Biswas et al., 2020). The numeric rating from online customer reviews can depict the product or personal reputation because customer experience can be reflected on the valence of electronic word-of-mouth (Chen et al., 2017). It can provide customers heuristic reference for their purchase decisions when making some judgments over product or service quality (Zhang et al., 2011). Moreover, evidence has shown that the badge system is the manifestation of seller reputation, which can bring successful booking behavior (Liang et al., 2017; Liang et al., 2020). In the accommodation-sharing economy, the information can be classified into user or customer-generated content (CGC) and marketer-generated content (MGC) (Biswas et al., 2020; Liang et al., 2020). According to the definition of CGC and MGC, online rating is one of the primary forms of CGC, while the badge systems or status profiles are MGC. Both of them constitute the effective reputation in P2P accommodations, which significantly influence the count of customer reviews (Biswas et al., 2020). Hereby, we define online rating as customer-generated reputation and the badge status (e.g., the "superhost" badge) as marketer-generated reputation.

4.3 Hypotheses Development

4.3.1 Effects of self-presentation on sales performance

In the online environment, self-presentation is a strategic way to project an attractive image because it manifests computer-mediated communication and makes receivers form a perception of senders (Peng, 2020). Prior research has found that self-presentation positively contributes to relationship management and perceived usefulness, although customers concern the privacy (Min & Kim, 2015). It is not only valuable in social network websites (Raban, 2009; Ostermaier-Grabow & Linek, 2019) but also in the e-commerce platform. For example, self-presentation can develop customer trust and promote the purchase behavior because information disclosure indicates that service providers have put effort into self-marketing (Zhang et al., 2020).

Self-presentation plays an essential role in affecting customer resonance, which influences their perception of the product image and promotes their purchase intentions (Shang et al., 2017). Through disclosing personal information, self-presentation can convey social-oriented content and reduce social distance (García et al., 2019). In the accommodation-sharing platform, it is risky for customers when only reading product-related information. To obtain more trust-building signals, customers are likely to browse provider profiles because service providers participate in customer experience (Xie et al., 2019). Accordingly, host profile has become a crucial source of information, depicting service attributes and individual personality (Wu et al., 2017). Especially in P2P accommodations, where customers can share a room with service providers, the digital impression from host self-presentation determines their purchase behavior (Liang et al., 2020). Numerous studies have found that self-presentation is a strategic marketing attractiveness for unknown products, which is also a way of host-guest interaction to build a customer-brand relationship (Tussyadiah & Park, 2018; García et al., 2019). It has been found that host description is a pivotal signal to reduce social distance and affect customer booking intentions (Tussyadiah & Park, 2018). Thus, we propose the following hypothesis that

Hypothesis 1: Self-presentation positively affects sales performance.

In particular, the semantic feature of self-presentation is prominent for customers to make a reasonable judgment over service quality because it can help customers deeply understand service providers (Zhang et al., 2018). The semantic feature implied in self-presentation can mirror individual characteristics and personality tendencies (Tussyadiah & Park, 2018). The tendency of social interaction in host self-presentation depicts the possibility of offering warm hospitality from providers (Prayag & Ozanne, 2018). Evidence has shown that customers attach great importance to host hospitality and social interaction in the accommodation-sharing economy (Tussyadiah & Zach, 2017; Belarmino et al., 2019). From this standpoint, we assume that social-oriented self-presentation is crucial in determining customer purchases in P2P accommodations.

Using text mining, scholars have found two patterns in host self-presentation: a well-traveled social enthusiast and an individual with a certain profession (Tussyadiah & Park, 2018). Empirical evidence has demonstrated that the former plays a more decisive role in influencing customer purchase than the latter (Tussyadiah & Park, 2018). In addition, social-oriented self-presentation is conducive to forming an intimate seller-buyer relationship, which helps reduce social distance and affect customer behavior (Lin et al., 2019). Considering that social-oriented self-presentation can make customers perceive the hospitality from service providers and feel close to service providers, we hypothesize that

Hypothesis 2: Social-oriented self-presentation positively affects sales performance.

4.3.2 Interactions of self-presentation with reputation on sales performance

In e-commerce, self-presentation can reflect the signals regarding service attributes, such as social interaction and warm hospitality (Tussyadiah & Park, 2018). As a trustworthy signal, reputation is a critical index of depicting overall quality judged by the public (Abrate & Viglia, 2019). Evidence has shown that online reputation brings a competitive advantage because it carries an external clue of service or product quality, forming customer trust (Ert et al., 2016; ter Huurne et al., 2017). Based on the above, self-presentation and online reputation are two trust-building signals that capture customers' attention when searching and comparing the products across stores. Drawing on information processing theory (Thomas & McDaniel Jr, 1990), these signals may affect each other in influencing customer purchase decisions.

Prior literature has found that online review rating as a kind of customer-generated reputation can positively affect customer purchases because it reflects peer evaluation (Floyd et al., 2014). In addition, scholars have echoed that online customer rating exerts a positive and significant influence on economic success (Gossling et al., 2018). Using empirical validation, scholars have also found that customer-generated reputation maximizes revenues because it helps reduce the uncertainty of transactions (Abrate & Viglia, 2019). In e-commerce, customer-generated reputation not only exerts a direct effect on sales performance but also plays a moderating role in influencing the impact of trust-building signals on sales performance (Wang et al., 2016). Evidence has shown that reputation strengthens the effect of online reviews embedded in the product description on sales performance (Wang et al., 2016). As an essential signal to reduce social distance, host description may work together with customer-generated reputation to influence customer purchases because they offer external clues of the quality for customers' judgment (Xie et al., 2019). In particular, social-oriented self-presentation is effective in developing customer trust (Tussyadiah & Park, 2018). Evidence has shown that social-oriented self-presentation plays a decisive role in influencing sales performance and revenue because the use of social words can build a close seller-buyer relationship (García et al., 2019; Han et al., 2019). From these standpoints, we argue that customer-generated reputation can positively moderate the effects of self-presentation and social orientation on sales performance. Therefore, we propose that

Hypothesis 3a: Customer-generated reputation strengthens the positive effect of self-presentation on sales performance.

Hypothesis 3b: Customer-generated reputation strengthens the positive effect of social orientation in self-presentation on sales performance.

In the online platform, customers resort to multiple signals to reduce uncertainties from information asymmetry (Ert et al., 2016; Xie et al., 2019). The marketer-generated reputation is as crucial as host profiles and listing attributes in swaying customer purchase decisions (Ert et al., 2016). In purchasing a product, customers are likely to compare product/provider-related attributes across stores. It has been found that reputation and trust are two critical elements in swaying customer purchases (Ert et al., 2016). In P2P accommodations, the marketer-generated reputation awarded by the third-party platform conveys trustworthy signals to customers because it reflects the overall quality of service or products (Abrate & Viglia, 2019). Previous literature has shown that the badge of "superhost" is positively associated with e-WOM (e.g., review volume and ratings), which affects customers' willingness to pay (Liang et al., 2017). It has been found that marketer-generated reputation is a pivotal element that moderates the effect of marketer or host-generated content on customer purchases (Biswas et al., 2020). As a kind of marketer/host-generated content, self-presentation can help reduce social distance and develop customer trust toward service providers (García et al., 2019). Indeed, social-oriented self-presentation conveys some helpful information to help customers judge whether they have similar interests with service providers and evaluate how they get along with service providers. Moreover, previous research has provided strong evidence supporting that customers are likely to choose service providers who express social-oriented self-presentation because social value is crucial in P2P accommodations (García et al., 2019). Based on the above, we argue that the positive effects of self-presentation and social orientation on sales performance are strengthened by marketer-generated reputation. Accordingly, we hypothesize that

Hypothesis 4a: Marketer-generated reputation strengthens the positive effect of self-presentation on sales performance.

Hypothesis 4b: Marketer-generated reputation strengthens the positive effect of social orientation in self-presentation on sales performance.

4.4 Data and Methodology

4.4.1 Data and sample

To test our hypotheses, we target Airbnb in the United States as the research object. As a global leading accommodation-sharing online platform, Airbnb owns 700 million listings covering 191 countries and regions[9]. An official report has predicted that the number of users using Airbnb in the United States is expected to reach 45.6 million by 2022[10]. In addition, Airbnb is widely adopted as the sample to investigate customer behavior in the sharing economy (Ert et al., 2016; Lin et al., 2019; Zhang, 2019).

Data was collected on March 2, 2017, covering four cities, including Los Angeles, New York, San Francisco, and New Orleans. After deleting the duplicate data, we obtained 49,898 listings and 35,920 hosts. In addition, 71.81 percent of listings (35,535) have a homepage for provider self-presentation. Overall, our sample consists of listings characteristics, host profiles, and customer reviews. In the online platform, service providers can selectively disclose their profiles, shown in Figure 4.1. Moreover, there are two kinds of traditional reputations in Airbnb, the badge of "superhost" and average rating on the listings page, illustrated in Figure 4.2.

[9] https://www.airbnb.cn/ (accessed on 14 August 2020)
[10] https://www.statista.com/statistics/346589/number-of-us-airbnb-users/ (accessed on 14 August 2020)

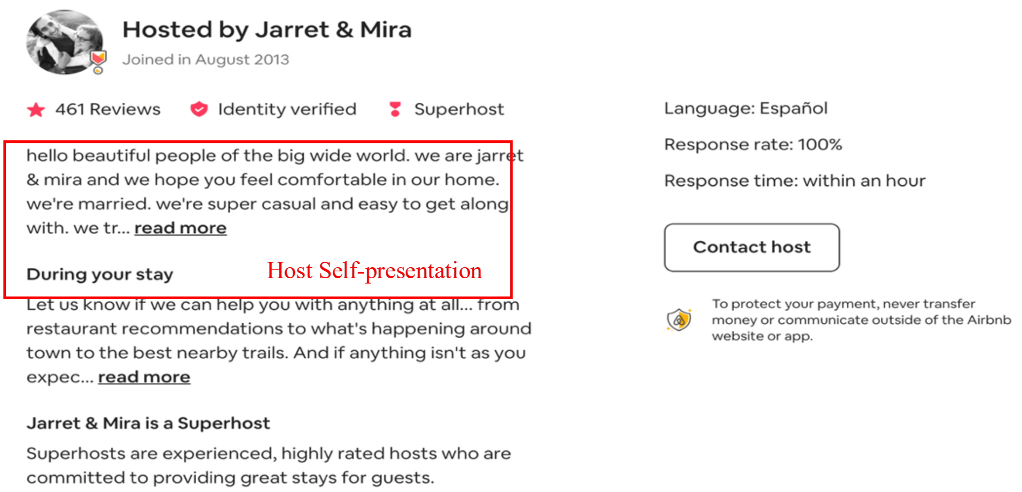


Figure 4.1 Host self-presentation in Airbnb

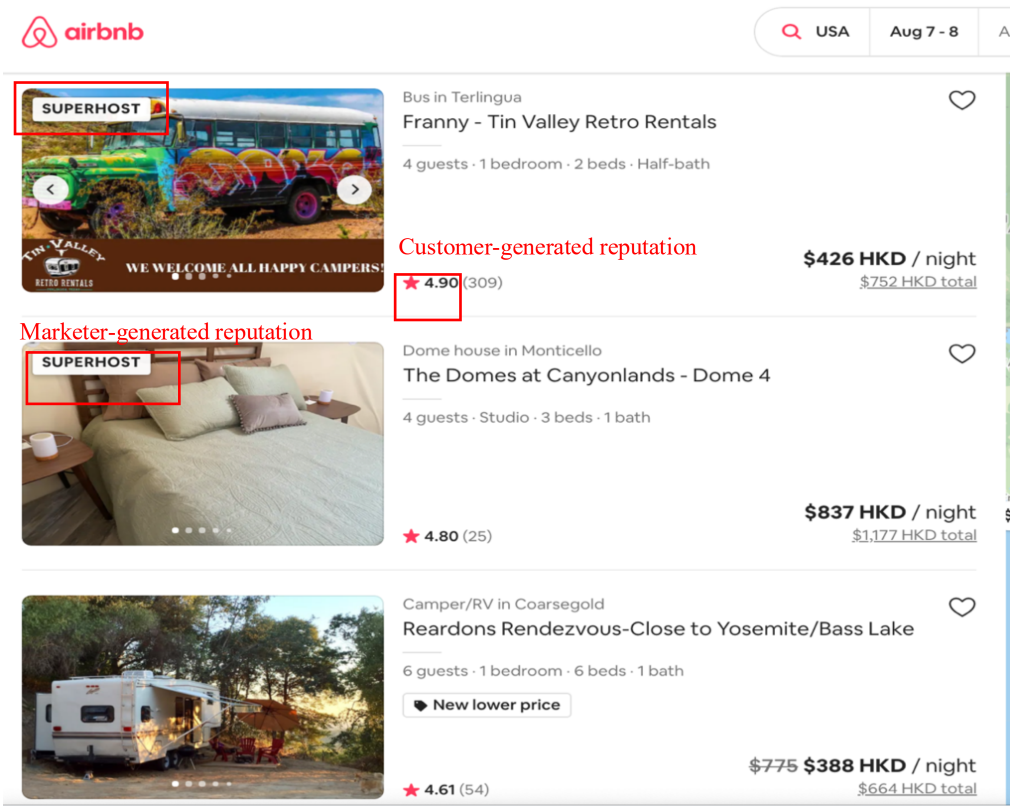


Figure 4.2 Reputation mechanisms in Airbnb

4.4.2 Text mining

(1) Topic extraction using LDA

Topic modeling is a research method of text mining to identify the general topics presented in the unstructured text (Berger et al., 2020). Latent Dirichlet allocation (LDA), a common tool of topic modeling, is an unsupervised machine learning algorithm to identify the hidden topics from large-scale documents (Blei et al., 2003). It is a topic-generation model containing three layers with words, topics, and documents, which is called a three-layer hierarchical Bayesian probability model (Blei et al., 2003).

According to the directed probability graph of topic modeling, as shown in Figure 4.3, the modeling process using LDA mainly includes two parts: one is to generate the distribution of topics θ for certain document and the topic z for certain words through inputting Dirichlet distribution α, the other is to generate the distribution of words φ for a certain topic and the words w through Dirichlet distribution β.

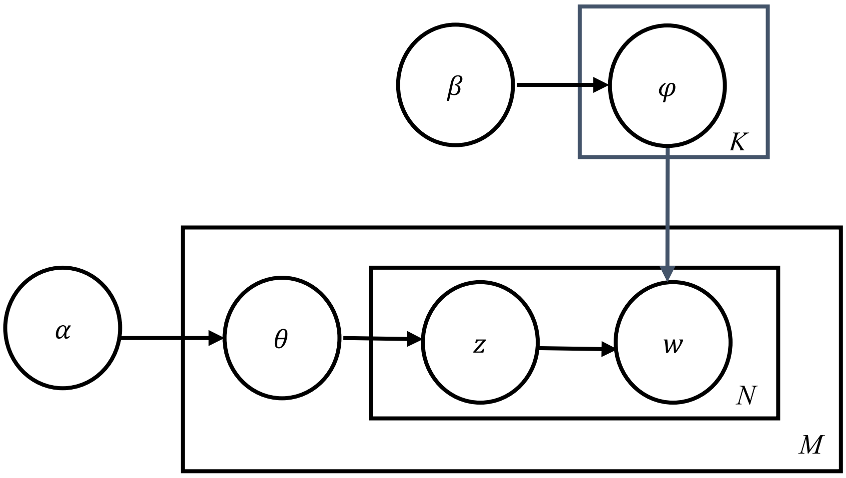


Figure 4.3 Directed probability graph of topic modeling

Before topic modeling, the optimal number of topics should be determined. Perplexity is a critical metric to evaluate the performance of the trained model and determine the optimal number of topics (Blei et al., 2003). In topic modeling, perplexity refers to that the trained model recognizes how many topics are contained in certain documents with uncertainty. Therefore, the lower the value of the perplexity, the smaller the uncertainty, and the better the clustering result. However, considering the operational cost, the turning point with the largest slope is selected as the optimal number of topics. The calculation equation of perplexity is as follows.

$$\text{perplexity(D)} = \exp\left(-\frac{\sum \log P(w)}{\sum_{d=1}^{M} N_d}\right)$$

Where N is the total number of words in the document, and P(w) refers to the probability of each word that occurs in the document. The calculation formula of P(w) is P(z|d) * P(w|z), among which P(z|d) represents the probability of certain topics in different documents, and P(w|z) represents the probability of certain words in various topics.

To calculate the perplexity, the primary step is to conduct data preprocessing. According to prior research (Berger et al., 2020), we adopt the following steps to complete data preprocessing. After obtaining the data, we clean the data by removing the nontextual and duplicate information. To eliminate useless words, we remove digits, punctuations, and nonmeaningful signs. Furthermore, we change all uppercase letters to lowercase letters to reduce the bias arising from spelling. Finally, we tokenize the text by eliminating common words from NLTK's list of English stopwords[11].

Next, we train the LDA model and identify the optimal topics implied in unstructured text. Following previous literature (Griffiths & Steyvers, 2004), the setting of α and β is critical to train a well-accepted LDA model. We set β=0.01 and takes α at three different levels, respectively 0.1, 1, and 50/K, and the number of topics ranges from 10 to 100. The result of perplexity is illustrated in Figure 4.4. We find that the optimal number of topics is 30, as the perplexity tends to be stable. In other words, the model generation ability is well-accepted when the number of topics is equal to 30.

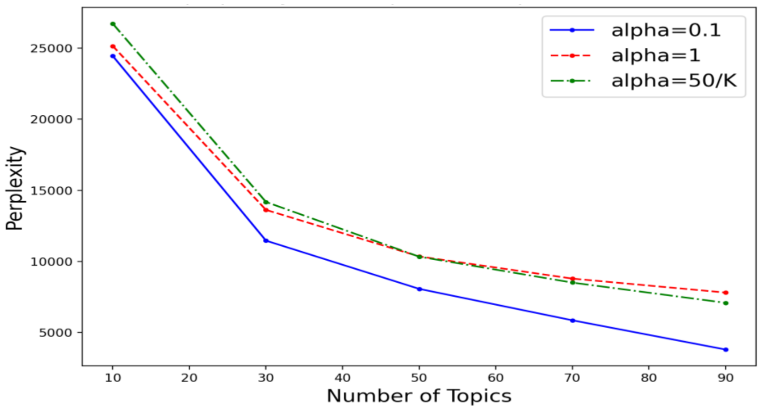


Figure 4.4 Perplexity under the different number of topics

After the optimal number of topics is determined, we conduct topic modeling to extract the general topics from all the samples. Table 4.1 presents the topics implied in the host self-presentation. According to previous literature (Tussyadiah & Park, 2018) and our research context, there are two dimensions of semantic feature, including social-oriented and official-oriented self-presentation. In terms of social-oriented self-presentation, hosts express their welcome to guests and introduce themselves as well-traveled social enthusiasts. For official-oriented self-presentation, hosts introduce basic information, such as profession, age, education, family, and living environment.

Table 4.1 Topics extracted from self-presentation

Dimensions	Categories	Topics	Feature words
Social-oriented	Welcome and hospitality	V2, V8, V10, V11, V15, V23, V30	Welcome, fun, life, nice, food, enjoying, coming, moment, days, friendly, happy, great, friend, enjoy
	Travel habits	V3, V16, V17, V20, V21, V24, V29	Travel, traveling, Airbnb, world, traveler, globe, vacation, retired

[11] https://gist.github.com/sebleier/554280 (accessed on 14 August 2020)

	Family	V4, V14, V26	Wife, girl, couple, mom, friends, family, married, woman,
Official-oriented	Work and education	V1, V7, V9, V13, V19, V25, V28	artist(s), actress, business, employed, writer, job, photographer, working, musician, professional, teacher, filmmaker, educator
	Age and details	V5	years, old
	Living environment	V6, V12, V18, V22, V27	Kitchen, neat, tidy, living, live, park, property

(2) Topic prediction based on machine learning

Machine learning is an approach that takes advantage of machines to recognize the human language, which can help process a large number of documents effectively (Berger et al., 2020). To identify whether self-presentation contains social orientation, we use machine learning algorithms to predict the topic of social orientation. First, we randomly select 1,106 samples and label them whether the self-presentation is social-oriented. The samples are independently annotated by two research students majoring in data science. The result is determined by another person who is an expert in online reviews when the label is inconsistent. Second, we construct a document-word matrix weighted by the word frequency, which is input into the machine. Third, we call various machine learning classifiers to train the model and select an optimal classifier. According to the classification performance of precision rate, recall rate, and F1 score, as well as micro-average, macro-average, and weighted average metrics (Schütze et al., 2008; Chang et al., 2020), illustrated in Figure 4.5, it is shown that the support vector machine (SVM) and logistic regression (LR) algorithms are better than other classifiers at identifying topics (micro-average:90%, macro-average:90%). To further assess the performance of binary classifiers, we calculate Area Under the Curve (AUC) and find that SVM is more appropriate for topic prediction (AUC_{SVM} =0.8981> AUC_{LR}=0.8961). Fourth, we employ SVM to predict the topic of social-oriented self-presentation. The results show that majority of self-presentation is social-oriented (59.37%).

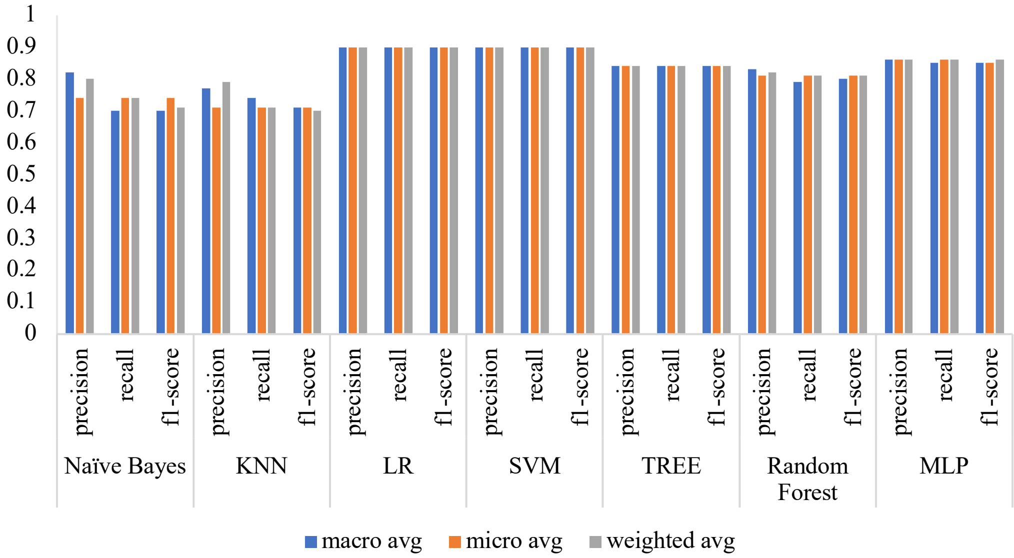


Figure 4.5 Classification performance for topic prediction

4.4.3 Model specification

To examine the effects of self-presentation and social orientation on sales performance, we consider a dummy variable about whether a host makes a self-presentation as the index of self-presentation. Furthermore, we capture its semantic feature about whether the self-presentation is social-oriented as the indicator of social orientation. Following previous literature (Ye et al., 2009; Ye et al., 2011), we use the number of reviews as the proxy of sales performance (Dellarocas et al., 2007). In terms of traditional reputations, we consider the badge of "superhost" as the index of marketer-generated reputation (Biswas et al., 2020) and take online rating as the indicator of customer-generated reputation (Gossling et al., 2018).

To rule out the interference of alternative explanations, we include a complete series of control variables in our proposed model. First, considering that host attributes may affect customer behavior (Wu et al., 2017), we add a variety of host signals, such as *Experience*, *Response rate*, *Number of listings*, *Picture*, and *Identity verification*, into the model. Second, previous research has demonstrated that listing characteristics play an essential role in influencing customer purchase decisions (Xie et al., 2019). We control some listing signals, such as *Price*, *Bathrooms*, *Bedrooms*, and *Beds*. Third, following previous literature (Biswas et al., 2020), we add living rules as control variables, such as *instant bookable*, *Cancellation policy*, and *Minimum of nights*. Finally, we control the *Genre* and *City* fixed effects in our proposed model. Table 4.2 describes the key variables in our model. In addition, Table 4.3 presents the correlations among variables. The absolute value of coefficients ranges from 0.001 to 0.694.

Moreover, we also conduct the variance inflation factor (VIF) test in the subsequent regression analyses. All the results indicate no severe multicollinearity of the model because the value of VIF is far less than 10.

Table 4.2 Variable description and statistics in Study 2

Variables	Description	Mean	Std. Dev	Min	Max
Sales performance	The total number of reviews that a listing receives	29.443	40.722	1	583
Self-presentation	Dummy variable about whether a host makes self-presentation	0.712	0.453	0	1
Social orientation	Dummy variable about whether host self-presentation is social-oriented	0.621	0.485	0	1
Superhost	Dummy variable about whether a host receives a badge of "superhost"	0.221	0.415	0	1
Review rating	Average rating that a listing receives	4.923	0.327	1	5
Experience	The time that a host starts to rent his/her listing	2013.735	1.758	2008	2017
Response rate	The number of host responses to the number of renter inquiries	0.940	0.159	0	1
Number of listings	The total number of listings that a host owns	4.573	23.066	0	1107
Picture	Dummy variable about whether a host offers the profile picture	0.998	0.043	0	1
Identity verification	Dummy variable about whether a host's identity is verified	0.754	0.431	0	1
Price	Average price of a listing	155.760	200.032	0	10000
Bathrooms	Number of bathrooms	1.241	0.610	0	8
Bedrooms	Number of bedrooms	1.283	0.864	0	10
Beds	Number of beds	1.808	1.374	1	16
Instant bookable	Dummy variable about whether a listing is instantly bookable	0.242	0.428	0	1
Cancellation policy	Cancellation policy on a scale of 1-3, 1=strict, 2=moderate, 3=flexible	1.658	0.776	1	3
Minimum of nights	The minimum of nights that a guest should book	3.308	6.719	1	365
Genre	Room type on a scale of 1-3, shared room=1, private room=2, and entire home=3	2.540	0.561	1	3
City	Categorial variable about the value of 1=San Francisco, 2=Los Angeles, 3=New Orleans, 4=New York	2.916	1.107	1	4

Table 4.3 Correlations in Study 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) Sales performance	1.000																	
(2) Self-presentation	0.200	1.000																
(3) Superhost	0.213	0.114	1.000															
(4) Review rating	0.124	0.079	0.114	1.000														
(5) Experience	-0.238	-0.375	-0.038	-0.074	1.000													
(6) Response rate	0.122	0.068	0.153	0.092	0.005	1.000												
(7) Number of listings	-0.037	0.028	-0.039	-0.038	0.006	0.023	1.000											
(8) Picture	0.010	0.027	0.008	0.024	-0.016	0.008	0.004	1.000										
(9) Identity	0.117	0.215	0.106	0.072	-0.243	0.058	0.017	0.071	1.000									
(10) Genre	0.036	0.020	0.039	0.060	-0.061	0.019	0.028	-0.005	0.043	1.000								
(11) Bathrooms	-0.057	0.006	0.006	-0.011	0.012	0.026	0.058	0.002	0.016	0.093	1.000							
(12) Bedrooms	-0.059	0.011	0.008	0.003	0.004	0.033	0.020	-0.004	0.025	0.253	0.576	1.000						
(13) Beds	-0.003	0.022	0.008	-0.008	0.028	0.049	0.052	-0.013	0.028	0.294	0.513	0.694	1.000					
(14) Price	-0.054	-0.006	0.015	0.009	-0.031	-0.008	0.026	-0.004	0.029	0.318	0.447	0.475	0.395	1.000				
(15) Minimum of	-0.069	0.039	-0.034	-0.027	-0.085	-0.016	0.010	-0.003	0.013	0.066	0.017	0.016	0.001	0.009	1.000			
(16) Instant bookable	0.093	-0.043	-0.019	-0.031	0.161	0.097	0.027	-0.011	-0.056	-0.053	-0.008	-0.020	0.042	-0.047	-0.054	1.000		
(17) Cancellation	-0.110	-0.108	-0.044	-0.034	0.100	-0.083	-0.067	-0.006	-0.094	-0.174	-0.091	-0.138	-0.170	-0.116	-0.058	-0.024	1.000	
(18) City	-0.109	-0.042	-0.154	-0.048	0.016	-0.090	-0.067	-0.013	-0.065	-0.077	-0.144	-0.082	-0.074	-0.075	-0.015	-0.029	-0.009	1.000

We employ the econometric models to estimate the effects of host self-presentation and social orientation for hosts who own self-presentation on sales performance and test the moderation effects of customer-generated and marketer-generated reputation on these relationships. Considering *Sales performance* whose standard deviation is much greater than the mean value (skewness=3.11), *Sales performance* with skewed distribution is taken log transformation (Trochim & Donnelly, 2001; Field, 2009). As a result, we construct the following models to address our research questions. The first equation estimates the effect of self-presentation and its interactions with reputation on sales performance. For hosts who own self-presentation, the second equation estimates the effect of social orientation and its interactions with reputation on sales performance. To make variables comparable, we standardize all the variables except for dependent variables.

$$\begin{aligned} Ln(Sales\ performance) &= \alpha_0 + \alpha_1\ Self - presentation + \alpha_2 Review\ rating \\ &+ \alpha_3 Superhost + \alpha_4 Self - presentation \times Review\ rating \\ &+ \alpha_5 Selfpresentation \times Superhost + \alpha_6 HOST + \alpha_7 LISTING \\ &+ \alpha_8 RULE + \alpha_9 City + \alpha_{10} Genre + \varepsilon \end{aligned}$$

$$\begin{aligned} Ln(Sales\ performance) &= \beta_0 + \beta_1 Social\ orientation + \beta_2 Review\ rating + \beta_3 Superhost \\ &+ \beta_4 Social\ orientation \times Review\ rating \\ &+ \beta_5 Social\ orientation \times Superhost + \beta_6 HOST + \beta_7 LISTING \\ &+ \beta_8 RULE + \beta_9 City + \beta_{10} Genre + \varepsilon' \end{aligned}$$

Where HOST represents a vector of host attributes controls that include *Experience, Number of listings, Response rate, Picture,* and *Identity verification*. LISTING is a vector that consists of listing characteristics, including *Price, Bedrooms, Bathrooms,* and *Beds*. RULE is a vector that represents accommodation rules, including *instant bookable, cancellation policy,* and *minimum of nights*. In addition, we include *City* and *Genre* controls as the proxy of the city of samples and the index of room type. ε and ε' are the random errors.

4.5 Results and Analyses

4.5.1 Estimations of provider self-presentation on sales performance

Table 4.4 illustrates the standardized estimations of provider self-presentation and its semantic feature as well as their interactions with reputation on sales performance.

Model 1 includes self-presentation, two kinds of reputations (i.e., superhost and review rating), and control variables, whereas Model 2 adds the two-way interactions of self-presentation with reputation. Model 3 includes social orientation, two kinds of reputations, and control variables, whereas Model 4 adds the two-way interaction of social orientation with reputation.

Models 1 and 2 show the estimations of sales performance by comparing hosts who make a self-presentation with those that don't introduce them. Evidence has shown that hosts who make efforts on self-presentation perform better in sales performance than those that don't introduce themselves (β=0.161, p<0.001). Thus, Hypothesis 1 is supported. It indicates that self-presentation can convey a trustworthy signal to reduce social distance and positively affect customer purchases, which is in line with prior literature supporting the importance of information disclosure in promoting sales (García et al., 2019).

Moreover, we find that social-oriented self-presentation positively influences customer purchases (β=0.043, p<0.001), shown in Model 4. Thus, Hypothesis 2 is supported. This finding is consistent with previous research arguing that social-oriented enthusiasm can help build a close buyer-seller relationship and develop customer trust (Tussyadiah & Park, 2018).

Table 4.4 Standardized estimations of sales performance in Study 2

D.V.: Ln (Sales performance)	Model 1	Model 2	Model 3	Model 4
Self-presentation	0.163***(0.006)	0.161***(0.006)		
Social Orientation			0.043***(0.007)	0.043***(0.007)
Superhost	0.283***(0.006)	0.287***(0.006)	0.275***(0.007)	0.276***(0.007)
Review rating	0.242***(0.005)	0.261***(0.006)	0.248***(0.007)	0.253***(0.007)
Self-presentation × Superhost		-0.035***(0.006)		
Self-presentation × Review rating		0.048***(0.005)		
Social orientation × Superhost				-0.012+(0.007)
Social orientation × Review rating				0.019**(0.006)
Experience	-0.194***(0.006)	-0.193***(0.006)	-0.203***(0.007)	-0.204***(0.007)
Response rate	0.107***(0.006)	0.107***(0.006)	0.112***(0.007)	0.112***(0.007)
Number of listings	-0.101***(0.005)	-0.101***(0.005)	-0.113***(0.007)	-0.113***(0.007)
Picture	-0.004(0.005)	-0.004(0.005)	-0.013+(0.007)	-0.013+(0.007)
Verified identity	0.063***(0.006)	0.063***(0.006)	0.049***(0.007)	0.050***(0.007)
Bathroom	-0.051***(0.007)	-0.051***(0.007)	-0.049***(0.009)	-0.050***(0.009)
Bedroom	-0.115***(0.008)	-0.115***(0.008)	-0.119***(0.010)	-0.119***(0.010)
Beds	0.135***(0.008)	0.134***(0.008)	0.138***(0.010)	0.139***(0.010)
Price	-0.109***(0.007)	-0.109***(0.007)	-0.120***(0.008)	-0.120***(0.008)
Minimum of nights	-0.117***(0.005)	-0.117***(0.005)	-0.122***(0.007)	-0.122***(0.007)
Instant bookable	0.148***(0.006)	0.148***(0.006)	0.160***(0.007)	0.160***(0.007)

Cancellation policy	-0.191***(0.006)	-0.190***(0.006)	-0.176***(0.007)	-0.176***(0.007)
City Dummies	Included	Included	Included	Included
Genre Dummies	Included	Included	Included	Included
Constant	2.467***(0.021)	2.466***(0.021)	2.544***(0.044)	2.545***(0.044)
Observations	49898	49898	35535	35535
VIF	1.56	1.53	2.29	2.18
F test	812.56***	745.72***	455.41***	414.62***
R-squared	0.2458	0.2475	0.2041	0.2044

Note: Standardized *b*'s are reported with robust standard errors in parentheses.
+p<0.1, *p < .05, **p < .01, ***p < .001

4.5.2 Interactions of provider self-presentations with reputation

Model 2 from Table 4.4 suggests that the customer-generated reputation (i.e., review rating) strengthens the positive effects of host self-presentation on sales performance (β=0.048, p<0.001). Furthermore, Model 4 from Table 4.4 illustrates the interaction effects of social-oriented self-presentation with the customer-generated reputation on sales performance. Specifically, the customer-generated reputation (i.e., review rating) strengthens the positive effect of social-orientated self-presentation on sales performance (β=0.019, p<0.01). Thus, hypotheses 3a and 3b are supported. These findings have demonstrated that online review rating, as a kind of customer-generated reputation, can work together with provider self-presentation to promote customer purchases. It confirms that trust-building signals from different sources can collectively determine customer purchase decisions (Choi et al., 2018). In addition, self-presentation depicts service attributes, while average rating mirrors product attributes. Both of them can mutually complement each other, collectively contributing to customer purchases (Lee et al., 2011)

On the other hand, Model 2 from Table 4.4 indicates that the marketer-generated reputation (i.e., superhost) weakens the positive effect of host self-presentation on sales performance (β=-0.035, p<0.001). Similarly, Evidence has shown that the positive effect of social-oriented self-presentation on sales performance is weakened by the marketer-generated reputation (i.e., superhost) (β=-0.012, p<0.1), as shown in Model 4. Contrary to our expectations, hypotheses 4a and 4b are not supported. One possible reason is that the badge of superhost and host self-presentation are marketer/seller-generated content. Both of them may suppress each other because they have the same and overlapping signals of reflecting host attributes (Xie & Lee, 2015). Prior literature has reported that the badge of superhost can mirror host responsiveness and reliability, which to some extent can be reflected on social-oriented self-presentation (Gunter, 2018). Figure 4.6 presents the moderating role of marketer-generated (i.e., superhost) and customer-generated reputations (i.e., review rating) in influencing the relationship between self-presentation and sales performance.

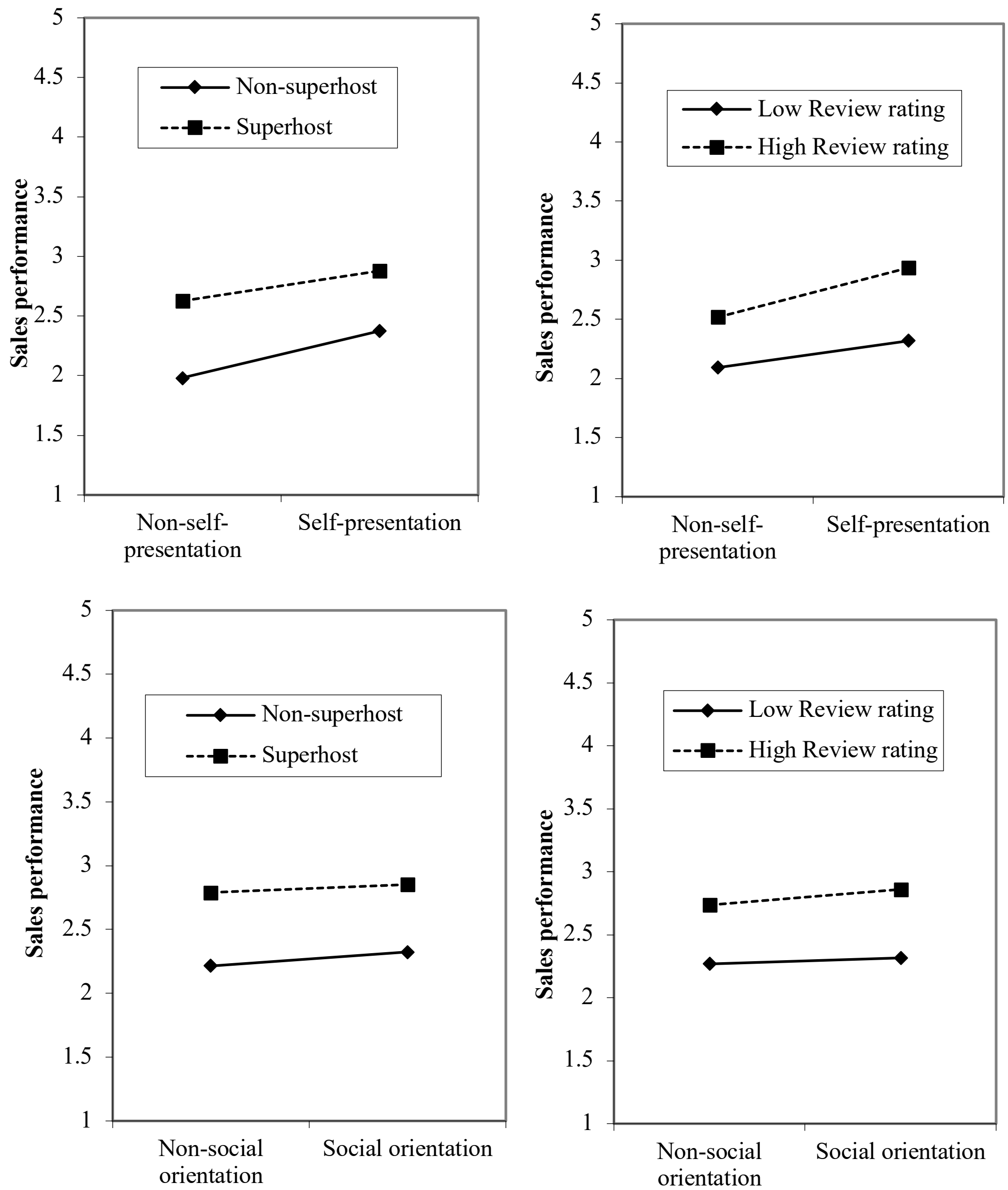


Figure 4.6 Moderating role of online reputation

4.5.3 Robustness check

We adopt multiple measures to ensure robustness, as shown in Table 4.5. First, we have included city and genre fixed effects in our proposed model. Second, due to that monthly review volume may affect sales performance, we add monthly review volume controlled in our proposed models, as shown in Models 1 and 2. Third, considering the pivotal role of textual length in self-presentation (García et al., 2019), we include *Length* as the control variable, as presented in Model 3. Fourth, we adopt the Poisson regression to repeat our proposed model where the dependent variable is not taken log transformation, as shown in Model 4 and Model 5. In summary, the results are in line with our previous findings.

Table 4.6 presents the results addressing the endogeneity of self-presentation and measure errors caused from text mining. Because online reputation may influence self-presentation in determining sales performance from Model 2 from Table 4.4, self-presentation is endogenous because it may be affected by online reputation. The coefficients of interaction terms between online reputation and self-presentation using OLS would be inflated. Following previous literature (Hamilton & Nickerson, 2003; Zhang et al., 2018), we use two-stage least square (2SLS) regressions to correct for the endogeneity. In the first stage, we regress self-presentation on customer-generated reputation, marketer-generated reputation, and control variables to obtain the residual, as illustrated in Model 6 from Table 4.6. Evidence has shown that customer-generated reputation (review rating: $\beta=0.034$, $p<0.001$) and marketer-generated reputation (superhost: $\beta=0.076$, $p<0.001$) positively affect self-presentation, supporting the use of the 2SLS regression to correct the endogeneity. In the second stage, we use the residual as the index of self-presentation and add interaction terms between the residual of self-presentation and online reputation (i.e., customer-generated reputation and marketer-generated reputation) to test the contingency effects, as shown in Model 7. In addition, considering the measurement errors from text mining (Yang et al., 2018), we introduce 5% errors through human intervention by changing the prediction results of social orientation to the opposite ones, illustrated in Model 8. Overall, the results are consistent with our findings, suggesting the robustness of our research framework.

Table 4.5 Robustness check for variable and model selection

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-presentation	0.118***(0.005)				0.172***(0.001)
Social Orientation		0.019***(0.005)	0.022**(0.007)	0.035***(0.001)	
Superhost	0.131***(0.005)	0.121***(0.005)	0.271***(0.007)	0.164***(0.001)	0.172***(0.001)
Review rating	0.169***(0.005)	0.168***(0.005)	0.252***(0.007)	0.236***(0.002)	0.261***(0.002)
Self-presentation × Superhost	-0.025***(0.005)				-0.021***(0.001)
Self-presentation × Review rating	0.039***(0.004)				0.023***(0.002)
Social orientation × Superhost		-0.012*(0.005)	-0.015*(0.007)	-0.009***(0.001)	
Social orientation × Review rating		0.011*(0.005)	0.018**(0.006)	0.005**(0.002)	
Experience	-0.313***(0.005)	-0.321***(0.005)	-0.202***(0.007)	-0.265***(0.001)	-0.262***(0.001)
Response rate	0.032***(0.004)	0.042***(0.005)	0.111***(0.007)	0.140***(0.001)	0.146***(0.001)

Number of listings	-0.059***(0.004)	-0.066***(0.005)	-0.118***(0.007)	-0.197***(0.003)	-0.162***(0.002)
Picture	-0.009*(0.004)	-0.011*(0.005)	-0.013*(0.007)	-0.003**(0.001)	0.002*(0.001)
Verified identity	0.039***(0.004)	0.023***(0.005)	0.047***(0.007)	0.035***(0.001)	0.047***(0.001)
Bathroom	-0.036***(0.006)	-0.043***(0.007)	-0.049***(0.008)	-0.040***(0.001)	-0.039***(0.001)
Bedroom	-0.018**(0.007)	-0.013+(0.008)	-0.119***(0.010)	-0.094***(0.001)	-0.089***(0.001)
Beds	0.049***(0.006)	0.048***(0.008)	0.135***(0.010)	0.094***(0.001)	0.094***(0.001)
Price	-0.043***(0.005)	-0.049***(0.006)	-0.119***(0.008)	-0.198***(0.002)	-0.216***(0.002)
Minimum of nights	-0.018***(0.004)	-0.020***(0.005)	-0.121***(0.007)	-0.209***(0.002)	-0.199***(0.001)
Instant bookable	-0.057***(0.004)	-0.050***(0.005)	0.160***(0.007)	0.143***(0.001)	0.143***(0.001)
Cancellation policy	-0.143***(0.004)	-0.142***(0.005)	-0.173***(0.007)	-0.111***(0.001)	-0.124***(0.001)
Monthly review volume	0.842***(0.005)	0.849***(0.006)			
Length			0.077***(0.007)		
City Dummies	Included	Included	Included	Included	Included
Genre Dummies	Included	Included	Included	Included	Included
Constant	2.521***(0.014)	2.436***(0.037)	2.533***(0.044)	3.329***(0.002)	3.114***(0.002)
Observations	49898	35533	35535	35535	49898
VIF	1.46	2.24	2.14	2.18	1.53
F test/LR chi2	2519.47***	1552.32***	403.51***	319178.7***	507731.4***
R-squared/Pseudo R-square	0.5375	0.5014	0.2072	0.1913	0.2346

Note: Standardized *b*'s are reported with robust standard errors in parentheses.
+p<0.1, *p < .05, **p < .01, ***p < .001

Table 4.6 Robustness check about the endogeneity and measurement errors

	Model 6	Model 7	Model 8
Self-presentation		0.147***(0.005)	
Social Orientation			0.022**(0.007)
Superhost	0.076***(0.004)	0.296***(0.006)	0.277***(0.007)
Review rating	0.034***(0.004)	0.248***(0.005)	0.253***(0.007)
Self-presentation × Superhost		-0.019**(0.006)	
Self-presentation × Review rating		0.021***(0.005)	
Social orientation × Superhost			-0.014*(0.007)
Social orientation × Review rating			0.017**(0.006)
Experience	-0.340***(0.004)	-0.248***(0.006)	-0.204***(0.007)
Response rate	0.039***(0.004)	0.114***(0.006)	0.113***(0.007)
Number of listings	0.028***(0.004)	-0.097***(0.005)	-0.114***(0.007)
Picture	0.012**(0.004)	-0.002(0.005)	-0.012+(0.007)
Verified identity	0.113***(0.004)	0.081***(0.006)	0.051***(0.007)
Bathroom	0.00002(0.005)	-0.051***(0.007)	-0.050***(0.009)
Bedroom	-0.012+(0.006)	-0.117***(0.008)	-0.119***(0.010)
Beds	0.049***(0.006)	0.143***(0.008)	0.139***(0.010)
Price	-0.032***(0.005)	-0.114***(0.007)	-0.121***(0.008)
Minimum of nights	0.012**(0.004)	-0.115***(0.005)	-0.122***(0.007)
Instant bookable	0.011**(0.004)	0.150***(0.006)	0.160***(0.007)
Cancellation policy	-0.056***(0.004)	-0.199***(0.006)	-0.177***(0.007)
City Dummies	Included	Included	Included
Genre Dummies	Included	Included	Included
Constant	0.029+(0.015)	2.471***(0.021)	2.549***(0.044)
Observations	49898	49898	35535

VIF	1.57	1.49	2.18
F	559.88***	740.19***	412.82***
R-squared	0.1758	0.2461	0.2037

Note: Standardized *b*'s are reported with robust standard errors in parentheses.
+p<0.1, *p < .05, **p < .01, ***p < .001

4.6 Discussions

4.6.1 Findings

Using the combination of text mining and econometric analysis, we examine the effects of self-presentation and social orientation as well as their interactions with two kinds of reputations on sales performance. Through topic extraction using LDA, we find two formats in the host self-presentation, social-oriented and official-oriented. This classification result is consistent with previous literature identifying two patterns of self-presentation: a well-traveled social enthusiast and an individual with a specific profession (Tussyadiah & Park, 2018). Furthermore, we select the SVM algorithm to identify whether host self-presentation is social-oriented and extract the variable of interest for subsequent econometric analysis.

For econometric analysis, we find that host self-presentation can positively affect sales performance. It supports prior literature arguing that information disclosure can develop customer trust and promote customer purchases (Xie et al., 2019). Moreover, we also find that social-orientated self-presentation positively influences sales performance. This finding suggests that the social-oriented expression evokes customer resonance and determines customer behavior (Han et al., 2019). In terms of the moderating role of online reputation, our study provides strong evidence supporting that the positive effects of self-presentation and social orientation embedded in self-presentation on sales performance are strengthened by customer-generated reputation (i.e., review rating). However, the marketer-generated reputation (i.e., superhost) weakens the positive impact of self-presentation and social orientation in self-presentation on sales performance. These findings not only verify that reputation is an essential factor in determining customer purchases (Wang et al., 2016) but also support a marked difference between customer-generated and marketer-generated reputations in moderating the relationship between provider self-presentation and sales performance (Abrate & Viglia, 2019). Notably, online reputation plays a decisive role in moderating the effects of trust-building signals on customer purchase decisions in the sharing economy (Ert et al., 2016). Moreover, it also indicates that multiple signals from the same sources may suppress each other because they release similar and overlapping signals that compete for consumers' attention. Otherwise, the signals from different sources may complement each other because they can reflect a variety of external clues of service or product quality (Xie et al., 2019).

4.6.2 Theoretical implications

Our study contributes to the existing literature on the sharing economy as well as service businesses. It offers new insights on marketing strategies by disclosing provider self-presentation and gaining the online reputation. First, our findings advance the social distance theory applied in the platform-mediated economy by emphasizing that self-presentation is an effective strategy of reducing social distance and developing customer trust. In particular, social-oriented self-presentation plays a crucial role in swaying customer purchases because social-oriented expression can help reduce social distance. These findings support that constructing a digital self can signal trustworthiness to customers when they feel the distance from service providers in the online environment (García et al., 2019; Xie et al., 2019). It not only confirms that self-presentation is a way to reduce social distance (Stephan et al., 2011) but also adds new insights into the communication literature by uncovering that social-oriented self-presentation enhances sales performance in e-commerce.

Second, our findings can provide new insights into the sharing economy literature by shedding light on the interactions of trust-building signals from customer-generated and provider/marketer-generated content in the online platforms. Specifically, supplier profiles are as crucial as online reputation in influencing customer purchases under the context of information asymmetry (Ert et al., 2016). More importantly, this study extends the signaling theory by illustrating that trust-building signals from different sources can complement each other and otherwise suppress each other. It has supported that trust-building signals may affect (i.e., suppress and complement) each other when they compete for customers' attention (Xie et al., 2019).

Third, our research proves that online impression management (e.g., host profiles and reputation) is an effective way to attract customers because it helps customers build a close seller-buyer relationship (Ellison et al., 2006; Liang et al., 2020). In particular, self-presentation and online reputation convey trust-building signals, which are pivotal in the sharing economy where service providers participate in customer experience (Tussyadiah & Park, 2018; García et al., 2019). In e-commerce, online reputation is a crucial factor in influencing customer purchase (Ert et al., 2016; Abrate & Viglia, 2019). This study can advance the existing knowledge by shedding light on the importance of online impression management by constructing a virtual image from two perspectives: self-presentation and online reputation.

Finally, we advance the existing knowledge on the online reputation by examining the crucial role of customer- and marketer-generated reputations in moderating the relationship between self-presentation and sales performance. Specifically, we have provided strong evidence supporting that customer-generated reputation complements, but marketer-generated reputation substitutes the positive effects of self-presentation and social orientation on sales performance. Accordingly, our findings uncover the marked difference between customer-generated and marketer-generated reputations in online platforms (Liang et al., 2020). It further confirms prior literature arguing that product reputation (i.e., review rating) differs from personal reputation (i.e., superhost) in moderating the effect of shared assets on achieved revenues in the sharing economy (Abrate & Viglia, 2019). Our research has supported the distinctive role of customer-generated reputation and marketer-generated reputation in moderating the relationship between trust-building signals and sales performance (Wang et al., 2016; Abrate & Viglia, 2019). Specifically, we enhance the in-depth understanding of online reputation by demonstrating the substitution effect of marketer-generated reputation and the complementary effect of customer-generated reputation between self-presentation and sales performance.

4.6.3 Managerial implications

Our findings provide some suggestions for service providers to optimize marketing strategies. Given the importance of self-presentation as an advertisement to attract potential customers, information disclosure is an appropriate choice to strengthen the seller-buyer relationship and reduce social distance in the context of information asymmetry. To excite customer resonance, it is suggested that service providers should disclose their personal information to introduce themselves by opening a homepage. More importantly, service providers should actively present their opinions and create a digital self via linguistic communication implied in their profiles to better reduce social distance in the online environment. For example, they can use informal language and express their hospitality/sociability by introducing their hobbies (e.g., family and travel) and showing the welcome to newcomers. To gain more popularity in sales, it is an effective way for service providers to obtain recognition and build an online reputation from the platform and customers. As for service providers who don’t own self-presentation, they should put more effort into obtaining marketer-generated reputation to attract customers. As for those who have a self-presentation, they should try their best to get a high average rating from customer reviews to create more benefits. On the other hand, as for service providers who have a low marketer-generated reputation, they should make a tremendous effort on their profiles to attract customers. In particular, service providers are suggested to create a good impression via self-presentation and electronic word-of-mouth because building an online image is crucial in developing customer trust and promoting customer purchase decisions.

Moreover, we also offer some ideas on managing the platform-mediated economy for market operators and platform designers. First, it is suggested that market operators should implement incentives to encourage service providers to create a virtual self by opening their homepage and introducing themselves. It can help sustain the sharing economy by attracting more customers to create greater benefits, which can be also beneficial to customers for seeking helpful information. Second, market operators should pay more attention to the setting of online reputation, putting an end to the review manipulation and formulating a transparent standard for marketer-generated reputation. For example, the assessment of the virtual badge or status should be made public, which can motivate service providers to optimize their service. Third, to improve customer experience in using the platform, platform designers should put host profiles and online reputation in an eye-catching position, which can help customers conveniently capture valuable information.

4.6.4 Limitations and future directions

Some limitations motivate our future studies. First, considering that host self-presentation remains unaltered. Thus, we cannot conduct a robust panel analysis. Second, we choose the number of reviews as the proxy of sales performance, as the information of listing sales is unavailable. Future research can gain valuable data on sales volume by cooperating with the associated business companies. Third, some psychological and linguistic features can be extracted from self-presentation, such as the use of the second-person pronoun, which may influence customer purchases. To comprehensively examine the effect of self-presentation, scholars in the future can obtain diverse and multiple signals from provider self-presentation, such as the sentiment and the expression style embedded in the unstructured text. Finally, we focus on Airbnb and cannot generalize results applicable to other platforms. For example, scholars can use data from Xiaozhu.com, a Chinese home-sharing online platform, to replicate our proposed model. In Xiaozhu.com, where service providers can selectively open personal webpages and disclose private information to the public, information disclosure can be considered as the indicator of host self-presentation. Besides, the number of disclosed items can, to some extent, measure the index of social orientation.

Chapter 5 Conclusions

5.1 Summary of Research Findings

In this thesis, we have conducted two interrelated studies to investigate the impact of CGC (i.e., review valence) and MGC (i.e., self-presentation) on sales performance and examine the boundary condition of these relationships. The major findings of these two studies are presented as follows.

In Study 1, integrating machine learning-based text mining with econometric models, we examine the effects of review valence (i.e., review sentiment and average rating) on sales performance and test the moderating role of seller popularity and property quantity in these relationships. Based on our analyses of 3,260 listings with customer reviews from Xiaozhu.com, a Chinese accommodation-sharing platform, we find that review valence (i.e., review sentiment and average rating) is positively associated with sales performance. Notably, we find no substantial difference between average rating and review sentiment in affecting sales performance. Moreover, seller popularity strengthens but property quantity weakens the positive effects of review valence on sales performance. Notably, the positive effect of review valence given to hosts characterized by popularity and personalization on sales performance is more prominent, and such hosts own fewer listings to achieve more historical sales.

In Study 2, we aim to empirically investigate the effects of host self-presentation and its semantic feature on sales performance and examine the moderating role of marketer-generated and customer-generated reputations in these relationships. Through a mixed-methods approach of latent Dirichlet allocation (LDA) and econometric models to deal with data from 49,898 listings from Airbnb, we find that host self-presentation is positively associated with sales performance. Specifically, social-orientated self-presentation positively affects sales performance. Moreover, we find that customer-generated reputation (i.e., review rating) strengthens, but marketer-generated reputation (i.e., superhost) weakens the positive effects of self-presentation and social orientation on sales performance.

5.2 Research Implications

5.2.1 Theoretical implications

This thesis contributes to the existing literature on e-commerce transactions by revealing what and how trust-building signals embedded in UGC affect sales performance. This thesis helps consolidate our understanding of the influence of UGC on choosing P2P accommodations. This thesis makes several theoretical contributions, as follows.

Study 1 provides empirical evidence to answer the debate over the relationship between review valence and sales performance. It confirms that review valence positively influences sales performance because it depicts collective wisdom and signals service or product quality (Floyd et al., 2014; You et al., 2015). In addition, our findings provide new insights by demonstrating the equal importance of review sentiment and average rating in affecting customer purchases. It can illustrate previous research that review sentiment better predicts average rating (Li et al., 2019; Zhu et al., 2020). More importantly, this study extends the signaling theory and observational learning theory by uncovering the complementary effect of seller popularity and the suppression effect of property quantity on the link between review valence and sales performance. It suggests that host profiles are as crucial as online customer reviews in carrying trust-building signals (Xu et al., 2021).

Study 2 enriches the extant knowledge by deeply understanding the effect of self-presentation on sales performance in P2P accommodations. It confirms that self-presentation is a way to reduce social distance by building an intimate seller-buyer relationship (García et al., 2019). Furthermore, our findings provide new insights into the existing literature on the online reputation by uncovering the substitution effect of marketer-generated reputation and the complementary effect of customer-generated reputation in the relationship between self-presentation and sales performance. One possible reason is that self-presentation can act as a virtual reputation built by service providers (Pera et al., 2016). Marketer-generated reputation has the same signal as self-presentation, and these two signals compete for customers' attention (Xie et al., 2019). Thus, it leads to the fact that the marketer-generated reputation substitutes the effect of self-presentation on sales performance.

5.2.2 Managerial implications

This thesis also has several implications for service providers to enhance their service operations management, as follows.

Study 1 provides the implications for service providers to optimize their service operations management in the sharing economy. First, it is suggested that service providers should build a harmonious relationship with customers, which can leave a good impression on the latter. Second, as for service providers who have a lower reputation, communicated by review valence, they are suggested to balance the tradeoff between listing quantity and service quality to create great benefits. For example, they can focus their energies on few listings to offer high-quality service or put more attention on managing multiple listings to improve their popularity. Finally, to achieve higher sales performance, service providers should take advantage of few listings to gain more sales and obtain recognized e-WOM by satisfying customer needs, which can help attract potential customers.

Study 2 offers managerial implications for service providers to formulate their marketing strategies through disclosing self-presentation and gaining online reputation. First, service providers are suggested to introduce themselves and build a virtual image by opening their homepages, which can reduce social distance and improve sales performance. Moreover, to incite customers' empathy, service providers should reveal their personality and hobbies in self-presentation. Second, in addition to building a personal reputation by making a self-presentation, service providers should pay more attention to improving product reputation, which can help them gain higher sales performance. Finally, service providers can also gain recognition by enhancing their comprehensive abilities if they are worry about the privacy issues of self-disclosure. In summary, service providers should put great effort into online reputation management, which is crucial in influencing customer purchase.

5.2.3 Policy implications

This thesis provides some suggestions for policy makers to regulate the sharing economy as well as service businesses. First, government agencies should notice the importance of online customer reviews. It is suggested that policy makers should put an end to fake reviews and formulate policies to restrict review manipulation. For example, they can develop some rules about review systems management into the market policy. Second, it is necessary to formulate rules and regulations to improve the ability assessment of service providers, which can energize and sustain the sharing economy. Considering that service attributes are as crucial as product features in determining customer purchases, policy makers are suggested to raise the barriers to entry for service providers and train them in professional ethics. Third, considering the importance of self-presentation, market regulators can help reduce the information asymmetry by implementing incentive mechanisms to motivate service providers to introduce themselves. It can not only help customers reduce perceived risks but also is beneficial for service providers to create profits. Fourth, to advance reputation systems, policy makers should set a transparent standard for assessing service providers and encourage customers to post any reviews about evaluating service or product quality.

5.3 Limitations and Future Research

This thesis is impossible without limitations, which pave the way for our future research. First, as the representatives of the accommodation-sharing economy, Xiaozhu and Airbnb have been selected to investigate customer behavior. Notwithstanding, this thesis lacks a systematic comparison of platform or cultural differences. In the future, it is promising that scholars make a contrast between Xiaozhu and Airbnb to summarize the unique features of various platforms and guests with different cultural backgrounds. Second, considering that service providers enter or exit more frequently, and cross-period user-generated content (i.e., review valence and self-presentation) for a listing is limited and insufficient for panel analysis, as shown in Study 1 and Study 2. It is suggested that scholars collect the long-term data for subsequent observations to repeat our proposed model using a panel analysis. In addition, the proposed model in this thesis can also be extended to other fields where the dataset is relatively large. Third, this thesis focuses on review valence and social-oriented content of self-presentation, ignoring some detailed information extracted from UGC, such as richness and readability. Considering the presence of multiple signals hidden in UGC, scholars can extract diverse and numerous information, such as the use of emotional words, to deeply gauge the predictors of sales performance. Fourth, because two studies in this thesis have adopted a mixed-methods approach of text mining and econometric models to link unstructured text and structured data to sales performance, there may exist measurement errors from variables generated by text analysis. It is necessary for scholars to improve the accuracy of topic or sentiment prediction by employing a refined algorithm, which can help reduce the measurement errors. Fifth, in this thesis, we consider the number of reviews as the proxy of sales performance because we cannot obtain the sales volume from online platforms. Future research can adopt alternative indicators, such as the ranking of recommendations and e-WOM, to measure sales performance. In addition, scholars can also cooperate with companies to acquire the actual sales data for each product. Moreover, due to the limitation of second-hand data, future research can introduce laboratory experiments and questionnaires to measure customers' purchase intentions as the index of sales performance by simulating the scene where customers read peer reviews and provider self-presentation. It can offer supplementary evidence to better understand UGC.

5.4 Concluding Remarks

With the advancement of social media, the sharing economy emphasizes customer experience and heavily relies on UGC. As two critical indicators of UGC, CGC and MGC play an essential role in swaying customer decisions. To provide prospective customers with heuristic reference, previous customers can evaluate their experience by posting reviews to either support or refute other reviewers. Moreover, service providers disclose their profiles and build a virtual image on online platforms. In the information asymmetry, customers are likely to seek trust-building signals from UGC and make their approach or avoidance in their initial choice set consideration. In this thesis, we examine the effects of customer reviews (i.e., review valence) and provider description (i.e., self-presentation) on sales performance. In addition, we also discuss the boundary condition of the relationship between UGC and sales performance. Using a combination of text mining and econometric model, this thesis links unstructured text and structured data to sales performance from the perspective of signaling theory.

This thesis deepens our understanding of signal-based decisions in the sharing economy as well as other service businesses by uncovering the importance of customer reviews and provider self-presentation in supporting customer decisions. On the one hand, Study 1 demonstrates that review sentiment is as crucial as average rating in influencing sales performance. More importantly, we also advance the knowledge by illustrating the complementary effect of seller popularity and the suppression effect of property quantity on the relationship between review valence and sales performance. The suggestions are helpful for managers to improve their service operations management by balancing their service effort between listing quantity and service quality. On the other hand, Study 2 shows that social-oriented self-presentation positively affects sales performance. Notably, this thesis adds new insights on the online reputation literature by revealing the complementary effect of customer-generated reputation (i.e., review rating) and the substitution effect of marketer-generated reputation (i.e., superhost) in the relationship between self-presentation and sales performance. Our findings support that online impression management is crucial for service providers to achieve profits, which provides practical suggestions on marketing strategies by introducing self-presentation and building a reputation. In summary, this thesis provides strong evidence supporting that multiple signals from online platforms will compete for customers' attention or work together to affect customer choices. The implications are helpful for service providers to attract customers by conveying trustworthy signals and for market operators to sustain the sharing economy by regulating the transparency of UGC. More importantly, it is suggested for government agencies to formulate policies to encourage information disclosure in a sustainable manner and put an end to malicious manipulation of UGC. We hope that this thesis can provide some insightful ideas on future research to deeply gauge the importance of UGC and understand customer behavior in e-commerce.

References

Abrate G.&Viglia G. (2019). Personal or product reputation? optimizing revenues in the sharing economy. *Journal of Travel Research*, 58(1), 136-148.

Ayeh J. K., et al. (2013). "Do we believe in tripadvisor?" Examining credibility perceptions and online travelers' attitude toward using user-generated content. *Journal of Travel Research*, 52(4), 437-452.

Bai X., et al. (2020). A note on the impact of daily deals on local retailers' online reputation: mediation effects of the consumer experience. *Information Systems Research*, 31(4), 1132-1143.

Bandura A.&McClelland D. C. (1977). *Social learning theory*. Prentice Hall: Englewood cliffs.

Banerjee A. V. (1992). A simple-model of herd behavior. *Quarterly Journal of Economics*, 107(3), 797-817.

Belarmino A., et al. (2019). Comparing guests' key attributes of peer-to-peer accommodations and hotels: mixed-methods approach. *Current Issues in Tourism*, 22(1), 1-7.

Belk R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research,* 67(8), 1595-1600.

Berger J., et al. (2020). Uniting the tribes: using text for marketing insight. *Journal of Marketing*, 84(1), 1-25.

Bertarelli S. (2015). On the efficacy of imperfect public-monitoring of seller reputation in e-commerce. *Electronic Commerce Research and Applications,* 14(2), 75-80.

Biswas B., et al. (2020). Examining the determinants of the count of customer reviews in peer-to-peer home-sharing platforms using clustering and count regression techniques. *Decision Support Systems,* 135, 113324.

Blei D. M., et al. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.

Bulchand-Gidumal J.&Melian-Gonzalez S. (2020). Why are ratings so high in the sharing economy? Evidence based on guest perspectives. *Current Issues in Tourism*, 23(10), 1248-1260.

Bulut Z. A.&Karabulut A. N. (2018). Examining the role of two aspects of eWOM in online repurchase intention: An integrated trust-loyalty perspective. *Journal of Consumer Behaviour,* 17(4), 407-417.

Chang W. L.&Wang J. Y. (2018). Mine is yours? Using sentiment analysis to explore the degree of risk in the sharing economy. *Electronic Commerce Research and Applications,* 28, 141-158.

Chang Y. C., et al. (2020). Using deep learning and visual analytics to explore hotel reviews and responses. *Tourism Management*, 80, 104129.

Chen K., et al. (2017). An influence framework on product word-of-mouth (WoM) measurement. *Information & Management,* 54(2), 228-240.

Chen L. R., et al. (2017). Who should pay for online reviews? Design of an online user feedback mechanism. *Electronic Commerce Research and Applications,* 23, 38-44.

Chen Y. F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior,* 24(5), 1977-1992.

Cheng M.&Jin X. (2019). What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*, 76, 58-70.

Cheng X. S., et al. (2019). An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust. *Tourism Management*, 71, 366-377.

Cheung C. M. K., et al. (2014). Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. *Decision Support Systems,* 65, 50-58.

Choi H. S., et al. (2018). The effect of intrinsic and extrinsic quality cues of digital video games on sales: An empirical investigation. *Decision Support Systems,* 106, 86-96.

Cohen J., et al. (2013). Applied multiple regression/correlation analysis for the behavioral sciences, Routledge.

Dellarocas C., et al. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing,* 21(4), 23-45.

Duan W., et al. (2009). Informational cascades and software adoption on the internet: an empirical investigation. MIS Quarterly, 23-48.

Duan W. J., et al. (2008). Do online reviews matter? - An empirical investigation of panel data. *Decision Support Systems*, 45(4), 1007-1016.

Ellison N., et al. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication,* 11(2), 415-441.

Ellway B. P. W. (2014). Is the quantity-quality trade-off in call centres a false dichotomy? Managing Service Quality.

Ert E.&Fleischer A. (2019). The evolution of trust in Airbnb: A case of home rental. *Annals of Tourism Research,* 75, 279-287.

Ert E., et al. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, 55, 62-73.

Field, A. (2009). Discovering statistics using SPSS:(and sex and drugs and rock’n’roll) (ed.). London, Royaume-Uni: Sage.

Floyd K., et al. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing,* 90(2), 217-232.

Fu S. X., et al. (2021). Staying in a hotel or peer-to-peer accommodation sharing? A discrete choice experiment with online reviews and discount strategies. *Internet Research,* 31(2), 654-676.

García M. N., et al. (2020). Be social! The impact of self-presentation on peer-to-peer accommodation revenue. *Journal of Travel Research*, 59(7), 1268-1281.

Gibbs J. L., et al. (2006). Self-presentation in online personals - The role of anticipated future interaction, self-disclosure, and perceived success in Internet dating. *Communication Research*, 33(2), 152-177.

Gossling S., et al. (2018). The manager's dilemma: a conceptualization of online review manipulation strategies. *Current Issues in Tourism,* 21(5), 484-503.

Griffiths T. L.&Steyvers M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5228-5235.

Gunter U. (2018). What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area. *Tourism Management,* 66, 26-37.

Hamari J., et al. (2016). The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology,* 67(9), 2047-2059.

Hamilton B. H.&Nickerson J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic organization*, 1(1), 51-78.

Han H., et al. (2019). Which appeals (ethos, pathos, logos) are the most important for Airbnb users to booking? *International Journal of Contemporary Hospitality Management*, 31(3), 1205-1223.

Hirata E. (2020). Service recovery and customer satisfaction in container liner shipping industry-an ordered logit approach. *International Journal of Shipping and Transport Logistics*, 12(6), 563-575.

Hua Y., et al. (2020). Monetary rewards, intrinsic motivators, and work engagement in the it-enabled sharing economy: A mixed-methods investigation of internet taxi drivers. *Decision Sciences,* 51(3), 755-785.

Ikkala T., et al. (2015). Monetizing network hospitality: Hospitality and sociability in the context of Airbnb. *In Proceedings of the 18th ACM conference on computer supported cooperative work & social computing,* 1033-1044.

Jia Y. L.&Liu I. L. B. (2018). Do consumers always follow "useful" reviews? The interaction effect of review valence and review usefulness on consumers' purchase decisions. *Journal of the Association for Information Science and Technology*, 69(11), 1304-1317.

Jin J. H., et al. (2015). Why users contribute knowledge to online communities: An empirical study of an online social Q&A community. *Information & Management*, 52(7), 840-849.

Josang A., et al. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems,* 43(2), 618-644.

Ju Y., et al. (2019). Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *International Journal of Hospitality Management*, 77, 342-352.

Kasabov E. (2016). Unknown, surprising, and economically significant: The realities of electronic word of mouth in Chinese social networking sites. *Journal of Business Research,* 69(2), 642-652.

Ke Q.&Acm (2017). Sharing means renting? An entire-marketplace analysis of Airbnb. In *Proceedings of the 2017 ACM on web science conference*, 131-139.

Kim K., et al. (2008). Effects of temporal and social distance on consumer evaluations. *Journal of Consumer Research,* 35(4), 706-713.

Krasnova H., et al. (2010). Online social networks: why we disclose. *Journal of Information Technology*, 25(2), 109-125.

Kwon S. (2020). Understanding user participation from the perspective of psychological ownership: The moderating role of social distance. *Computers in Human Behavior*, 105, 106207.

Lai K.-h., et al. (2020). Risk in port logistics: risk classification and mitigation framework. *International Journal of Shipping and Transport Logistics*, 12(6), 576-596.

Lee M. K. O., et al. (2011). Consumer's decision to shop online: The moderating role of positive informational social influence. *Information & Management,* 48(6), 185-191.

Li J., et al. (2016). Pros vs joes: Agent pricing behavior in the sharing economy. *Ross School of Business Paper*, 1298.

Li X. L., et al. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management,* 56(2), 172-184.

Liang S., et al. (2017). Be a "Superhost": The importance of badge systems for peer-to-peer rental accommodations. *Tourism Management,* 60, 454-465.

Liang S., et al. (2020). The importance of marketer-generated content to peer-to-peer property rental platforms: Evidence from Airbnb. *International Journal of Hospitality Management,* 84, 102329.

Liang S., et al. (2021). Tit for tat: understanding the responding behavior of property hosts on peer-to-peer rental platforms. *International Journal of Contemporary Hospitality Management*, 33(3), 1105-1126.

Lim W. M., et al. (2021). Home sharing in marketing and tourism at a tipping point: What do we know, how do we know, and where should we be heading? *Journal of Business Research,* 122, 534-566.

Lin C.-Y., et al. (2020). They support, so we talk: the effects of other users on self-disclosure on social networking sites. *Information Technology & People*.

Lin H. X., et al. (2019). Impact of tourist-to-tourist interaction on tourism experience: The mediating role of cohesion and intimacy. *Annals of Tourism Research*, 76, 153-167.

Lin P. M. C., et al. (2019). Spend less and experience more: Understanding tourists' social contact in the Airbnb context. *International Journal of Hospitality Management,* 83, 65-73.

Luo Y.&Tang R. (2019). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *International Journal of Hospitality Management*, 80, 144-154.

Ma X., et al. (2017). Self-disclosure and perceived trustworthiness of Airbnb host profiles. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing,* 2397-2409.

Mariani M. M.&Borghi M. (2020). Online review helpfulness and firms' financial performance: An empirical study in a service industry. *International Journal of Electronic Commerce,* 24(4), 421-449.

Mauri A. G., et al. (2018). Humanize your business. The role of personal reputation in the sharing economy. *International Journal of Hospitality Management*, 73, 36-43.

Min J.&Kim B. (2015). How are people enticed to disclose personal information despite privacy concerns in social network sites? The calculus between benefit and cost. Journal of the Association for Information Science and Technology 66(4), 839-857.

Moon H., et al. (2019). Peer-to-peer interactions: Perspectives of Airbnb guests and hosts. *International Journal of Hospitality Management*, 77, 405-414.

Ostermaier-Grabow A.&Linek S. B. (2019). Communication and self-presentation behavior on academic social networking sites: An exploratory case study on profiles and discussion threads on ResearchGate. *Journal of the Association for Information Science and Technology,* 70(10), 1153-1164.

Park S. W.&Nicolau J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research,* 50, 67-83.

Peng K. (2020). To be attractive or to be authentic? How two competing motivations influence self-presentation in online dating. *Internet Research*, 30(4), 1143-1165.

Pera R., et al. (2016). Who Am I? How Compelling Self-storytelling Builds Digital Personal Reputation. *Journal of Interactive Marketing,* 35, 44-55.

Pitt C., et al. (2021). Accommodation eWOM in the sharing economy: automated text comparisons from a large sample. *Journal of Hospitality Marketing & Management,* 30(2), 258-275.

Prayag G.&Ozanne L. K. (2018). A systematic review of peer-to-peer (P2P) accommodation sharing research from 2010 to 2016: progress and prospects from the multi-level perspective. *Journal of Hospitality Marketing & Management,* 27(6), 649-678.

Raban D. R. (2009). Self-Presentation and the value of information in Q&A websites. *Journal of the American Society for Information Science and Technology*, 60(12), 2465-2473.

Raguseo E.&Vitari C. (2017). The Effect of Brand on the Impact of e-WOM on Hotels' Financial Performance. *International Journal of Electronic Commerce,* 21(2), 249-269.

Ren J., et al. (2021). Effect of crowd wisdom on pricing in the asset-based sharing platform: An attribute substitution perspective. *International Journal of Hospitality Management,* 94, 102874.

Roschk H.&Gelbrich K. (2017). Compensation revisited: A social resource theory perspective on offering a monetary resource after a service failure. *Journal of Service Research*, 20(4), 393-408.

Santos G., et al. (2020). Neutrality may matter: sentiment analysis in reviews of Airbnb, Booking, and Couchsurfing in Brazil and USA. *Social Network Analysis and Mining*, 10(1), 1-13.

Schütze, H., Manning, C. D., & Raghavan, P. (2008). Introduction to information retrieval. *Cambridge: Cambridge University Press*, 39, 234-265.

Shang S. S. C., et al. (2017). Generating consumer resonance for purchase intention on social network sites. *Computers in Human Behavior,* 69, 18-28.

Spence M. J. A. E. R. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review*, 92(3), 434-459.

Stephan E., et al. (2011). The effects of time perspective and level of construal on social distance. *Journal of Experimental Social Psychology*, 47(2), 397-402.

ter Huurne M., et al. (2017). Antecedents of trust in the sharing economy: A systematic review. *Journal of Consumer Behaviour,* 16(6), 485-498.

Thomas J. B.&McDaniel Jr R. R. (1990). Interpreting strategic issues: Effects of strategy and the information-processing structure of top management teams. *Academy of Management Journal,* 33(2), 286-306.

Thomsen C.&Jeong M. (2020). An analysis of Airbnb online reviews: user experience in 16 U.S. cities. *Journal of Hospitality and Tourism Technology,* 12(1), 97-111.

Trochim W. M.&Donnelly J. P. (2001). Research methods knowledge base (Vol. 2). Atomic Dog Pub.

Trope Y.&Liberman N. (2010). Construal-Level theory of psychological distance. *Psychological Review,* 117(2), 440-463.

Tussyadiah I. P. (2016). Strategic self-presentation in the sharing economy: Implications for host branding. *Information and Communication Technologies in Touris,* Springer: 695-708.

Tussyadiah I. P.&Zach F. (2017). Identifying salient attributes of peer-to-peer accommodation experience. *Journal of Travel & Tourism Marketing*, 34(5), 636-652.

Tussyadiah S. P.&Park S. (2018). When guests trust hosts for their words: Host description and trust in sharing economy. *Tourism Management,* 67, 261-272.

Viglia G., et al. (2014). Please, talk about it! When hotel popularity boosts preferences. *International Journal of Hospitality Management*, 42, 155-164.

Wang Z. S., et al. (2016). Saliency effects of online reviews embedded in the description on sales: Moderating role of reputation. *Decision Support Systems*, 87, 50-58.

Wu J., et al. (2021). How to increase customer repeated bookings in the short-term room rental market? A large-scale granular data investigation. *Decision Support Systems*: 143, 113495.

Wu J., et al. (2017). In sharing economy we trust: the effects of host attributes on short-term rental purchases. *International Journal of Contemporary Hospitality Management,* 29(11), 2962-2976.

Xie K.&Lee Y.-J. (2015). Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model. *Journal of Management Information Systems,* 32(2), 204-238.

Xie K. R., et al. (2019). Learning from peers: The effect of sales history disclosure on peer-to-peer short-term rental purchases. *International Journal of Hospitality Management*, 76, 173-183.

Xie K. R.&Mao Z. X. (2017). The impacts of quality and quantity attributes of Airbnb hosts on listing performance. *International Journal of Contemporary Hospitality Management,* 29(9), 2240-2260.

Xie K. R.&Mao Z. X. (2019). Locational strategy of professional hosts: effect on perceived quality and revenue performance of Airbnb listings. *Journal of Hospitality & Tourism Research,* 43(6), 919-929.

Xu X. (2020). How do consumers in the sharing economy value sharing? Evidence from online reviews. *Decision Support Systems*, 128, 113162.

Xu X., et al. (2021). The impact of information disclosure on consumer purchase behavior on sharing economy platform Airbnb. *International Journal of Production Economics,* 231, 107846.

Xu X. Y., et al. (2017). The impact of informational incentives and social influence on consumer behavior during Alibaba's online shopping carnival. *Computers in Human Behavior,* 76, 245-254.

Yang M., et al. (2018). Mind the gap: Accounting for measurement error and misclassification in variables generated via data mining. *Information Systems*

Research, 29(1), 4-24.

Yao B., et al. (2019). Standing out from the crowd - an exploration of signal attributes of Airbnb listings. *International Journal of Contemporary Hospitality Management*, 31(12), 4520-4542.

Ye H., et al. (2019). Enhancing mobile data services performance via online reviews. *Information Systems Frontiers*, 21(2), 441-452.

Ye Q., et al. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management,* 28(1), 180-182.

Ye Q., et al. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, 27(2), 634-639.

Yi B., et al. (2021). Do the flipped impacts of hotels matter to the popularity of Airbnb? *International Journal of Contemporary Hospitality Management*.

You Y., et al. (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing,* 79(2), 19-39.

Zamani E. D., et al. (2019). Trust in the sharing economy: the AirBnB case. *Industrial Management & Data Systems,* 119(9), 1947-1968.

Zervas G., et al. (2020). A first look at online reputation on Airbnb, where every stay is above average. *Marketing Letters*, 32(1), 1-16.

Zhang C., et al. (2018). Contract learning in the aftermath of exchange disruptions: An empirical study of renewing interfirm relationships. *Industrial Marketing Management,* 71, 215-226.

Zhang J. R. (2019). What's yours is mine: exploring customer voice on Airbnb using text-mining approaches. *Journal of Consumer Marketing,* 36(5), 655-665.

Zhang L., et al. (2018). A computational framework for understanding antecedents of guests' perceived trust towards hosts on Airbnb. *Decision Support Systems*, 115, 105-116.

Zhang L., et al. (2020). A text analytics framework for understanding the relationships among host self-description, trust perception and purchase behavior on Airbnb. *Decision Support Systems*, 133, 113288.

Zhang S. S., et al. (2019). Does more accessibility lead to more disclosure? Exploring the influence of information accessibility on self-disclosure in online social networks. *Information Technology & People,* 32(3), 754-780.

Zhang Z. Q., et al. (2011). Determinants of hotel room price An exploration of travelers' hierarchy of accommodation needs. *International Journal of Contemporary Hospitality Management*, 23(6-7), 972-981.

Zhao X. Y., et al. (2015). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management,* 27(6), 1343-1364.

Zhu L., et al. (2020). Sentiment and guest satisfaction with peer-to-peer accommodation: When are online ratings more trustworthy? *International Journal of Hospitality Management,* 86, 102369.