

Research Paper

The 7 Ps marketing mix of home-sharing services: Mining travelers' online reviews on Airbnb

Linchi Kwok^{a,*}, Yingying Tang^b, Bei Yu^c^a The Collins College of Hospitality Management, California State Polytechnic University Pomona (Cal Poly Pomona), Pomona, CA, USA^b Whitman School of Management, Syracuse University, Syracuse, NY, USA^c School of Information Studies, Syracuse University, Syracuse, NY, USA

ARTICLE INFO

Keywords:

Marketing mix
Machine learning
Big data analytics
Online reviews
Home-sharing
Airbnb

ABSTRACT

The 7 Ps model is a very useful tool in helping service firms solve managerial issues in marketing. Guided by the 7 Ps marketing mix framework, a big-data, supervised machine learning analysis was performed with 1,148,062 English reviews of 37,092 Airbnb listings in San Francisco and New York City. The results disclose similar patterns in both markets, where travelers shared their experience about Service Product and Physical Evidence most often; Price and Promotion were the least mentioned elements. Furthermore, through a series of comparisons of Airbnb's 7 Ps marketing mix among the listings managed by different types of hosts, multi-unit and single-unit hosts seem to offer similar services with a small observable difference; whereas superhosts and the ordinary hosts deliver different services. This study makes valuable methodological contributions and provides practical marketing insights for hoteliers and the hosts and webmasters on home-sharing websites. Policymakers should pay special attention to multi-unit hosts.

1. Introduction

The marketing mix framework, traditionally known as 4 Ps (Product, Promotion, Price, and Place), is a dominant marketing management theory and a widely-used managerial tool that helps identify the principal components of a product (Fan et al., 2015). According to Goi's (2009) review of the marketing mix, the development of the 4 Ps Model can be chased back to the 60 s when Borden (1964) discussed the marketing mix concept. Borden (1964) proposed a marketing mix framework with 12 elements, including product planning, pricing, branding, channels of distribution, personal selling, advertising, promotions, packaging, display, servicing, physical handling, as well as fact findings and analysis, aiming to support teaching in marketing, problem-solving in business, and a systematic way of thinking about marketing. McCarthy (1964) extended Borden's (1964) marketing mix concept by introducing the 4 Ps model. Although the 4 Ps model also received some criticisms (e.g., Constantinides, 2006), the framework remains to be a very useful tool in helping business to solve the managerial issues regarding market segmentation, marketing positioning, and differentiation, allowing firms to focus on one or more elements of the marketing mix (Möller, 2006).

Later, Booms and Bitner (1981) extended the 4 Ps model by adding three Ps, including Participant (or People), Physical Evidence, and

Process for service products. As compared to 4 Ps, the 7 Ps model provides a more comprehensive and detailed framework in analyzing a service product's marketing mix (Loo and Leung, 2018). Besides its ability to assess the critical marketing elements of a service product, the 7 Ps Model can also be applied to other areas as a generic marketing framework (Rafiq and Ahmed, 1995) in such inquiries as qualitative content analysis (e.g., Loo and Leung, 2018), survey studies (e.g., Harrington et al., 2017), and big data analytics for business/marketing intelligence (e.g., Fan et al., 2015).

In this study, we adopted the 7 Ps marketing mix framework to guide our big-data, supervised machine learning analysis of an immense quantity of online review information that is created by consumers in the cyber marketplace. Through the analysis of online reviews, researchers can discover new meaningful and often real-time business insights, compliment the traditional data sources from surveys that are primarily static (Xiang et al., 2015). Generally, consumers perceived online reviews, also known as an essential form of user-generated content (UGC), more helpful and credible than the information provided by the service providers (Mauri et al., 2018). Travelers profoundly rely on the information revealed on online reviews for travel planning (Kwok and Xie, 2016). Current literature has reported online reviews' possible impacts on consumer satisfaction, consumer experience, and such business performance of RevPAR (revenue per available

* Corresponding author at: The Collins College of Hospitality Management, Cal Poly Pomona, 3801 West Temple Avenue, Pomona, CA 91768 USA.

E-mail address: lkwok@cpp.edu (L. Kwok).

room) for hotels, price, and market share (Kwok et al., 2017). Examining consumers' online reviews as a critical information source for business insights through the lens of the 7 Ps marketing mix framework can possibly provide valuable insights for relevant businesses.

We put our research setting on the revolutionary player in the lodging industry — Airbnb as the leading platform for home-sharing business. Since Airbnb's inception in 2008, the company has experienced exceptional growth. If measured by the number of rooms or listings in the market, Airbnb had already doubled the size of the world's largest hotel company (Xie and Kwok, 2017). Airbnb made disruptive impacts on the lodging industry (Blal et al., 2018; Heo and Blengini, 2019; Guttentag, 2015; Karlsson et al., 2017). Most recently, major hotel chains also see the value of the home-sharing market; they began entering the market to combat the competition from Airbnb and other home-sharing or vacation rental operators (El-Bawab, 2019). While a large number of research studies have assessed travelers' experience, motivations, and profiles with a cross-sectional survey or mixed methods (e.g., Jiang et al., 2019; Ju et al., 2019), there is a lack of systematic analysis or a clear understanding about the marketing mix of a home-sharing product/service.

Hence, the purpose of this study is two-folded. First and foremost, we aim to discover the essential elements of home-sharing products' marketing mix that are mentioned or appreciated the most by travelers, allowing the practitioners to draw insightful business intelligence. Second, we are aiming to join the scholarly discussion regarding big-data analytics for business intelligence (e.g., Fan et al., 2015; Zhao et al., 2019; Xiang et al., 2015) by providing an example where new managerial insights can be revealed through a systematic analysis of the colossal amount of the publicly-accessible, textual data created by consumers. Besides this study's methodological implications, the results of the 7Ps analysis can assist the hosts and webmasters on home-sharing websites and hoteliers to develop the differentiation and positioning strategies for the products they manage. Policymakers may also be able to find additional empirical evidence from this study to support their decisions on what type of restrictions are needed to regulate the home-sharing business in the lodging market.

As what is followed, we firstly introduce the research background of this study with a brief review of the relevant literature, including (a) home-sharing services as an interruptive player in the lodging industry, (b) the unique characteristics associated with the online reviews posted on home-sharing websites like Airbnb, and (c) the rationale of why additional analysis is needed to compare the marketing mix among the Airbnb listings that are managed by different types of service providers (i.e., different types of hosts). We then summarize our discussion in a visual diagram that conceptualizes our work with three specific research questions that guide our big-data analysis. After we describe our newly-developed data-analytic approach in detail, we present the results according to the order of the three research questions. In the end, we highlight this study's implications.

2. The Research Background and Relevant Literature

2.1. Home-sharing websites: The interruptive player in the lodging industry

Home-sharing websites provide a cyber marketplace where travelers can find alternative accommodation options from the residents living in a tourist destination other than hotels or hostels. Not only do home-sharing websites compete directly with hotels in the market by fulfilling the accommodation needs for travelers (Xie et al., 2020b), the business model of home-sharing websites also made disruptive impacts on the lodging industry.

Deviated from the traditional hotel development or asset management approach, home-sharing websites adopted a new business model for growth. For instance, they broke into the lodging market with a unique selling proposition with the offerings that hotels and hostels were not yet familiar with (Karlsson et al., 2017), allowing them to

spark enormous demands from travelers. Meanwhile, home-sharing websites provide a marketplace for people who want to make extra incomes from the under-utilized space they possessed as "micro-entrepreneurs," fueling the supply of the home-sharing facilities to the market (Abrate and Viglia, 2019; Kwok and Xie, 2019).

Home-sharing websites' additional disruptive impact on the lodging industry comes from its two-way reservation process¹ (Kwok and Xie, 2018). Hotels and hostels usually accept a reservation as soon as they receive a request with a valid form of payment. The reservation process on a home-sharing website also begins with a traveler browsing the available listings in a tourist destination. Once the traveler identifies an ideal, however, s/he must submit a request to the host who manages the listing. Then, it may take a few hours or even days for the host to evaluate the traveler's request before accepting or rejecting the request. Such reciprocal selection process is built upon the mutual trust between a traveler and a host (Yang et al., 2019), where the buyer and the seller develop the initial trust of the other party based on the "rich" information attached to the buyer's or seller's profile (Cui et al., 2017). As a result, home-sharing websites not only urge all travelers and hosts to share the rich demographic information about themselves in the cyber marketplace (Kwok and Xie, 2018) but also encourage them to evaluate each other upon the completion of a stay in the form of online reviews.

Generally speaking, online reviews have already become a significant, influential information source that affects consumers' purchasing decisions across a variety of products/services (e.g., Ahani et al., 2019; Padma and Ahn, 2020; Zhao et al., 2019). The online reviews about a home-sharing facility could play an even more essential role in influencing other travelers' purchasing decisions than those reviews about a branded hotel or hostel because listings on home-sharing websites do not have any brand affiliations, as what a hotel or hostel would usually have. According to the signaling theory, buyers' uncertainty or information asymmetry about a seller or a product can be reduced if certain information clues about the seller are presented (Connelly et al., 2011). In this case, a hotel's brand affiliation can serve as a critical information clue for travelers as they make a purchasing decision for a hotel stay. When travelers browse the options on a home-sharing website, however, there is no "brand affiliation" as an indicator of what services a listing or a host would offer; they must rely on the online reviews about a host and a listing for their decision-making.

2.2. Online reviews on home-sharing websites as a source of business intelligence

The online reviews shared by travelers is an essential component of UGC (Xiang et al., 2015). Consumers perceived UGC as a more helpful and credible information source than the information cues provided by the service providers (Mauri et al., 2018). UGC has made transformational changes in marketing practices because companies can no longer fully control what messages they want consumers to hear about their products or brands (Fader and Winer, 2012). At the same time, the colossal quantity of information in UGC also presents a fantastic opportunity for businesses to "listen" to their consumers and gain better insights about their customers as well as the competitors (Netzer et al., 2012). Through an analysis of the rich content in UGC, hospitality and tourism companies can draw critical business intelligence (BI) for decision makings (Talón-Ballesteros et al., 2018). In today's data-driven business environments, BI from big data analytics has become more than just an add-on utility; companies must rely on BI to sustain its

¹ As a means to reduce bias and potential discriminations against certain groups of travelers, Airbnb recently introduced the "instant book" option for hosts, enabling them to accept a request for a stay without the reviewing process if they choose to, the practice that is normally adopted by hotels or hostels (Ting, 2017). Nevertheless, the two-way reservation process remains to be a common practice at this point for home-sharing services.

competitive advantages in the market (Mariani et al., 2018).

Despite the critical role that online reviews play in consumers' purchasing decisions on a home-sharing stay, there is only scattered literature that reports relevant BI through the analysis of the rich textual data presented in online reviews or other forms of UGC. Kwok et al. (2017), for example, conducted a systematic analysis of 67 online review studies published in seven hospitality and tourism journals. They reported that online reviews could have a significant influence on consumer decision making regarding trust and attitude toward a brand/product, booking intention, satisfaction, and consumer experience, in addition to such business outcomes as hotel RevPAR, price, and market share. Methodology wise, while qualitative studies/reviews, surveys, and experiments were the commonly-used methods adopted in online review studies, a sub-stream of research using the quantitative big-data analytic approach also emerged in the literature (Kwok et al., 2017). In the home-sharing setting, Wiles and Crawford (2017) conducted a qualitative content analysis of 910 reviews of 48 hosts/properties according to the four aspects of the experience economy (i.e., education, esthetics, entertainment, and escapism), from which they found education dimension was the most represented content. Luo and Tang (2019) conducted a modified latent aspect rating analysis to better understand the hidden dimensions in the textual reviews on Airbnb. Their results suggest that consumers often comment on five aspects of the Airbnb services, including communication, experience, location, product/service, and value, where joy and surprise are the primary emotions identified. Cheng and Jin (2019) took a quantitative, unsupervised machine learning (Leximancer) and sentiment analysis approach to examine 170,124 online reviews on Airbnb in the Sydney, Australia market. Their results reveal that "location," "amenities, and "host" are the most mentioned attributes, and there is a strong positive bias in travelers' comments (e.g., a likelihood score of 66 % for positive sentiments vs. 1% for negative sentiments where "host" is mentioned).

Botsman (2017) attributed such a review inflation issue in the peer-to-peer economy, where reviewers are overwhelmingly positive, to the reciprocal selection process because the two-way review mechanism in the cyber marketplace adds social pressure on both the service providers and receivers, pushing them to write nice comments about the other party. In another in-depth manual content analysis, Bridges and Vásquez (2018) also found that travelers often expressed their negative experience in seemingly positive reviews with nuanced and subtle cues, making many online reviews on Airbnb look "alike" as they all sound very positive at a glance. Accordingly, the traditional sentiment analysis approach or the method of counting the vocabulary may not be ideal in examining the immense yet overwhelmingly positive online-review data on Airbnb. In this study, we report a unique data mining approach, where we adopted the 7 Ps Model as the guiding framework in analyzing the online reviews created by the travelers who stayed in an Airbnb facility. Through the lens of the 7 Ps Model, we assessed travelers' overall experience with the home-sharing products/services and, more importantly, compared and contrasted their experience with different types of service providers (hosts).

2.3. The marketing mix of home-sharing products/services: The three research questions

The 7 Ps marketing mix framework can be useful in helping marketers make decisions regarding segmentation, positioning, and differentiation (Möller, 2006). Even for the same type of products with different brands (e.g., shampoos of different brands), marketers may still want to take different actions to improve a product's marketing mix for better sales outcomes (Hanssens et al., 2014). Airbnb is a new product type that makes a disruptive impact on the lodging industry (Guttentag, 2015), an analysis of Airbnb's marketing mix will provide useful BI for the marketers working in hotel and home-sharing sectors as they make informed business decisions. Accordingly, our first interest is to address the following research question:

RQ1. What does the marketing mix revealed from consumers' online review data on Airbnb inform us about travelers' experience on home-sharing products?

As suggested in the signaling theory, consumers' uncertainty or information asymmetry in a business transaction about a product or a service provider can be significantly reduced if the service provider can present certain information clues to the buyers (Connelly et al., 2011). Unlike the regular hotel products where brand affiliations are perceived as an important information clue for the product's service quality, travelers rely on a listing's product attributes (e.g., safety and photos) as well as a host's personal reputation (e.g., being a superhost or not) for a purchasing decision (Abrate and Viglia, 2019). Recent research studies reported that a host's personal reputation could have an even more significant impact than a listing's product attributes on a home-sharing product's performance, such as the popularity of a listing (Mauri et al., 2018) and a listing's monthly revenue (Abrate and Viglia, 2019). It is hence not surprising to see that home-sharing websites publish a large amount of information about a host to reduce travelers' uncertainty or information asymmetry during the purchasing process. For example, a host can carry labels of a multi-unit or single-unit host, depending on how many units a host manages or how many connected accounts/listings a host operates, and of a superhost. Being a multi-unit or single-unit host, or being a superhost, could be a significant information clue for travelers as they assess the service quality of a home-sharing facility.

For example, the idea of sharing underutilized space among friends and family members is not new (Xie et al., 2020b). The original intents of many home-sharing websites were to build a cyber marketplace to connect those consumers who want to earn some extra income from the under-utilized space they possess with the travelers who need a space to stay (Guttentag, 2015). Very soon, the home-sharing marketplace attracted a large number of individuals who want to harvest more profits as the entrepreneurial service providers (Xie et al., 2019). More hosts are now turning their part-time, short-term residential rental business into a full-time job as a semi-professional micro-entrepreneur by managing more than just one unit (Kwok and Xie, 2018; Mauri et al., 2018). When a host is managing more than one listing (i.e., a multi-unit host), they must divide their attention to different groups of travelers; whereas they may devote their full attention to one group of travelers if they would have only managed one listing (i.e., a single-unit host). The lack of attention from a multi-unit host might result in a lower level of interactions between the host and the travelers. Or, they must be more proactive in identifying more efficient ways of handling customer-service issues, such as implementing more automatic services. Multi-unit hosts and single-unit hosts might end up offering a different experience to travelers, which will very likely reflect on the marketing mix of the services they provided. More importantly, multi-unit hosts may create even more significant threats to hoteliers because they can achieve higher revenues in the market through the effective use of pricing strategies (e.g., Gibbs et al., 2018; Kwok and Xie, 2018; Magno et al., 2018) and differentiated operational strategies (e.g., Xie et al., 2020b). Along this line of research, Mauri et al. (2018) also found that the number of connected accounts (i.e., the number of units that a host manages) has a significant negative impact on a listing's popularity as measured in rating, the number of reviews, times saved to a traveler's wish lists, and the likelihood of the host being categorized as a superhost.

When a host or a hotel/hostel manager wants to draw more specific implications in marketing positioning and differentiation, or when a policymaker wants to identify the significant impact from multi-unit vs. single-unit hosts, it becomes critical to also assess travelers' experience on a home-sharing facility based on the services provided by the multi-unit vs. single-unit hosts. Consequently, this study also aims to answer:

RQ2. Do travelers comment on the similar element(s) of the marketing mix for their home-sharing stays with the multi-unit hosts and the single-unit hosts? In other words, do travelers share similar experiences

for their home-sharing stays with the multi-unit hosts and the single-unit hosts through a comparison of the two product types' marketing mix?

Recently, the dominant home-sharing website, Airbnb introduced the "superhost"² category to label the "top-rated," "most experienced" hosts. According to Airbnb (2020), a service provider must meet the following criteria to be qualified as a superhost: (a) s/he has a 4.8 or higher average overall rating from at least 50 % of the guests being hosted in the past year; (b) s/he has hosted 10 or more stays in the past year or 100 nights over at least three stays; (c) no cancellations occurred in the past year except for extenuating circumstances; and (d) s/he respond to 90 % of new messages within 24 hours. As indicated in the signaling theory (Connelly et al., 2011), the superhost status provides a critical information clue for being a good indicator of the exceptional and reliable services provided by the host. The truth is having the superhost qualification can significantly increase a listing's reputation (Mauri et al., 2018) and monthly revenue (Abrate and Viglia, 2019). Hence, we propose that the marketing mix of the services provided by superhosts vs. the ordinary hosts (non-superhosts) could also be different, leading to the third and the final research question of this investigation:

RQ3. Do travelers comment on the similar element(s) of the marketing mix for their home-sharing stays with the superhosts and the ordinary hosts? In other words, do travelers share similar experiences for their home-sharing stays with the superhosts and the ordinary hosts through a comparison of the two product types' marketing mix?

As a summary, Fig. 1 provides a conceptual diagram of our analysis. Our first interest was to identify the marketing mix of home-sharing products in general, according to the 7 Ps Model through supervised machine learning with the immense textual data from the online reviews on Airbnb. Then, our second interest was to assess if the marketing mix differs according to the services provided by various types of hosts, including multi-unit vs. single-unit hosts and superhosts vs. the ordinary hosts.

3. Methodology

We adopted a five-step process in our data analysis. First, we collected and cleaned the textual online review data on Airbnb. Second, we prepared a sample dataset and manually annotated the 7 P dimensions in each sentence of the online reviews. Third, we used the manually-annotated results/data and machine learning algorithms to train a 7 P prediction model. Fourth, we tested and validated our machine learning model with a separate dataset. Fifth, we used the validated algorithms to analyze a large dataset to answer the three research questions.

3.1. Data collection and cleaning

We downloaded the data from InsideAirbnb.com, an independent, non-profit website that provides publicly accessible data collected from Airbnb.com. InsideAirbnb.com periodically downloads the data from Airbnb.com, including all the listings in selective markets, all reviews under each listing, and the unique characteristics of each listing (e.g., the entire unit or with shared space, managed by a superhost or not, managed by a host with one or more listings, and etc.). We selected two top Airbnb markets in the U.S.: San Francisco (SFO), a gateway city in the West Coast and where Airbnb began its business, and New York City (NYC), a gateway city in the East Coast and one of the top tourist destinations. We downloaded 233,070 reviews of 4,381 Airbnb listings

from the SFO market in September 2018 and 1,047,337 reviews of 32,985 Airbnb listings from the NYC market in October 2018. We used the SFO sample in training and tuning the machine learning algorithms and the NYC sample to validate the algorithms before applying the finalized algorithms to analyzing the full dataset to answer the research questions.

We broke down the reviews into individual sentences using the NLTK (The Natural Language Toolkit) sentence splitter under the assumption that people usually expressed one idea in one sentence, which in this case was one P or one element of the 7 Ps Model. With the effort of removing the non-English reviews from our datasets, we tested three stopword methods with a random sample of 200 sentences (such as in similar studies of Fong et al., 2006 and Yin and Yao, 2008), including Google Cloud Translation API (Google, 2019), the langid.py package (Lui and Baldwin, 2012), and the Guess Language (2019) package. Based on the accuracy of these programs and their availability for the usage in a large dataset like ours, we adopted the Guess Language package to clean the entire dataset, which achieved 99 % accuracy in our test. In the end, we identified 94.3 % (219,833) English for 4,376 listings and 5.7 % (13,288) non-English reviews in the SFO market, whereas we found 88.6 % (928,229) English for 32,716 listings and 11.4 % (119,023) non-English reviews in the NYC market.

3.2. Data annotation

To avoid poor operationalization where a sentence could be mistakenly coded in multiple categories of the 7 Ps model, we adjusted the definitions for each of the 7 P elements that are more specific for Airbnb listings (as in Table 1). For example, we recognized Physical Evidence, such as the bedding, the furniture, and kitchen equipment, could also be a critical component of an Airbnb Product. We then emphasized Product as the "Service Product" in this study, which is defined as the element for Airbnb as the "words that describe the overall impression of the intangible, experiential product." We defined the Physical Evidence of Airbnb as "words that describe the tangible aspect of the experiential product, such as the physical attributes and the facility of a listing."

Because the 7 Ps Model had not yet been applied to the analysis of Airbnb reviews, we conducted a preliminary coding on 10 reviews (132 sentences) that were randomly selected from the SFO sample to test (a) if the 7 Ps Model was suitable for annotating the Airbnb reviews and (b) if a sentence was the "right" unit for annotation. This approach is similar to those in natural language processing with machine learning methods (e.g., Pustejovsky and Stubbs, 2012). Our results suggested the proposed 7 Ps Model in Table 1 fit well with the annotation except that the model did not cover one type of sentences where travelers shared their advice about the neighborhood or their personal information. For example, travelers mentioned: "We had appointments all day Tuesday at UCSF;" "We searched for two months for a place to stay in SF;" and "I recommend Balboa Teriyaki, Americana and Purple Kow." We thus added an extra element to the 7 Ps Model as "Travelers" (TR) to include this type of sentences/information and finalized our coding schema as "7 Ps + TR," as outlined in Table 1.

Additionally, our results also confirmed that sentence was a suitable unit for coding. We only found 3% of the sentences mentioned two or more aspects in the 7 Ps Model. For instance, the sentence of "Lovely host with a beautiful location." mentioned both PP and PL. Likewise, the sentence of "Good location, nice and clean, easy to checkin." could be coded as PL, PE, and PS. Because these sentences might "confuse" the machine learning algorithms, and they only accounted for a tiny portion of the data, we decided to let the annotators to assign more than one element to a sentence but then remove these multi-element sentences from the training set. Using this approach also meant our prediction model would only label one element for each sentence. To further evaluate the reliability of the coding schema, we used Cohen's Kappa to assess the inter-coder agreement, with a result of 0.86, indicating a near-to-perfect agreement.

² Please visit <https://www.airbnb.com/superhost> for more details about the "superhost" program.

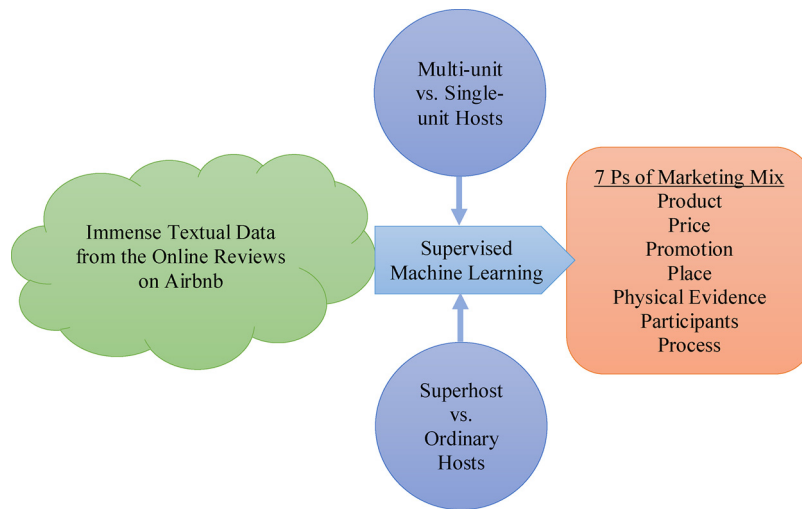


Fig. 1. The conceptual diagram of the proposed analysis.

Table 1
The 7 Ps + TR annotation schema.

Element	Definition	Examples
Service Product (PT)	Words that describe the overall impression of the intangible, experiential product.	<i>Would recommend 100 %.</i> <i>A disappointed experience.</i> <i>A great experience.</i> <i>It felt like home.</i>
Price (PR)	Words that indicate the price or value of the experiential product.	<i>If you like to cook your own meals and need to save some \$\$, you're in luck here.</i> <i>Rather than renting two hotel rooms, we split this 2-bed Airbnb and probably saved \$300 per night.</i>
Place (PL)	Words referring to the location of a listing.	<i>Short walk to the Bart station.</i> <i>Location is perfect – safe and pleasant neighbourhood closed to public transportation.</i>
Promotion (PO)	Words comparing what the traveler(s) observed against a listing's photos or descriptions on the website.	<i>The place was exactly as advertised!</i> <i>It's much more beautiful in person, indeed a hidden gem in downtown SF.</i> <i>The listing description was very accurate.</i>
Participants (PP)	Words mentioning the host(s) or people/pets in the listing.	<i>The listing is pretty close to the images posted.</i> <i>The hosts were also extremely friendly and accommodating.</i>
Physical Evidence (PE)	Words that describe the tangible aspect of the experiential product, such as the physical attributes and the facility of a listing.	<i>[host] was a great host-very helpful and accommodating.</i> <i>Super comfy bed, plenty of space to spread out, all kitchen necessities available, and right in the heart of the mission.</i> <i>Details like flowers, wine openers, britta water pitcher, shampoo, added to our relaxation.</i> <i>The room was spacious, quiet, and comfortable.</i>
Service Process (PS)	Words that emphasize the process where the traveler(s) received a service, with or without the interactions with the host.	<i>From the moment we first contacted the hosts to arriving they were very accommodating, making us feel very welcome upon arrival and taking time to show us around the room and giving great tips about the local area.</i> <i>[host name] left detailed instructions for us upon our arrival.</i>
Traveler (TR)	Words that are irrelevant to the experiential product (Airbnb) itself.	<i>She also helped me in before hand to find out the best transportation for me from the airport.</i> <i>We had ice cream for dessert two night in a row (Bi-Rite! The best!) and walked home with our cones feeling very happy indeed.</i> <i>I travel to San Francisco frequently, and was tired of staying in union square, financial district, etc.</i>

We then built a training dataset with 1000 randomly selected reviews from the SFO market, which included 4281 sentences, aligning with the approaches in natural language processing with machine learning methods recommended in Pustejovsky and Stubbs (2012). The two authors who participated in the reliability test of the coding schema earlier each coded half of this training dataset. Eventually, we coded 4120 sentences with one element and 161 (3.8 %) sentences with more than one element. Therefore, we retained the 4120 sentences being coded with only one element as the training dataset to avoid confusion. For the rest of the experiments, the most probable element was used if the test dataset contained examples that might be labeled with more than one element.

3.3. Model training

We employed a two-step process in training the model: sentence

vectorization and model building. During sentence vectorization, each sentence was transformed into a word vector, with each word as a dimension and its frequency or presence as a value. We evaluated a few common options for vectorization, including tokenization, lemmatization, stopwords elimination, term frequency, and TFIDF (term frequency-inverse document frequency) for choosing the best option in training the prediction model.

The word vectors were then sent to machine learning algorithms for building the prediction models, using SVM (support-vector network), the most comment text classification algorithm. We also chose the linear kernel for our experiment because most text classification problems are linearly separable (Joachims, 1998).

We used 5-fold cross-validation as an evaluation method to choose the best vectorization option and to tune the SVM algorithm parameters. We interpreted the model performance with the F-1 measure and confusion matrix, where we adopted the macro-averaging F-1 score

Table 2
The most informative features for prediction.

PT	PE	PL	PS	PP	PR	PO	TR
trip	wifi	downtown	process	met	spent	do	share
would	comfortable	bus	clear	hospitable	should	internet	rented
recommend	decor	bart	checking	advice	for	listing	wish
perfect	cozy	public	instructions	welcoming	spending	exactly	and
highly	beds	uber	arrival	providing	worth	as	now
experience	bed	walk	quick	kind	cost	reviews	nights
again	view	located	communication	helpful	cleaning	photos	going

that gives each element equal weights when averaging the F-1 scores from all categories because the eight categories in our training dataset were not evenly distributed. Macro-averaging F-1 score is a common evaluation measure that has been widely used in studies using SVM (Feng et al., 2020; Wang et al., 2015). We tested six prediction models, including “base (unigram Boolean),” “base + remove punctuation and digits,” “base + remove stopwords,” “based + lemmatization,” “based + bigram,” “TFIDF unigram,” and “TFIDF unigram + bigram,” all with a high macro-averaging F-1 score (≥ 0.81). We ultimately chose the TFIDF unigram model as the final model as it has the highest macro-averaging F-1 score.

Because SVM is a linear classifier, where the trained model assigns a weight to each feature as an indicator of the importance of feature prediction, the linguistic patterns that the model learned from the training data become the most important feature. Table 2 shows the top 10 features for each element of the “7 Ps + TR,” which are indeed commonly used in the corresponding categories and thus further validate the trained model.

3.4. Model generalization

To test the generalizability of our prediction model trained with the SFO data, we performed a generalization test with a different dataset from NYC. The best prediction model was used to analyze the entire NYC dataset, and the prediction result from a random sample of 200 reviews was further evaluated manually for accuracy. We manually annotated the sentences and then compared the human annotations against the machine predictions. The macro F-1 scores for the SFO and NYC samples were comparable to the cross-validation results on the training data reported in Section 3.3. We concluded that the model could be generalized across different datasets from the SFO and NYC markets.

3.5. Data analysis to answer the three research questions

We then applied the validated algorithms to analyze the English reviews in the SFO and NYC markets individually because the number of reviews from the NYC market is significantly higher than the number in the SFO market. To a large extent, reporting the results from the SFO market, as well as the NYC market, also allowed us to double-check the prediction reliability of our algorithms.

4. Results and Discussion

4.1. The marketing mix of home-sharing (Airbnb) listings in general

To answer RQ1, we included in our analysis the entire 219,933 English reviews of 4,376 listings in the SFO market and 928,229 English reviews of 32,716 listings in the NYC market. Figs. 2 and 3 show the distributions of the 7 Ps marketing mix (plus “Travelers”) regarding travelers’ experience of their Airbnb stays in the SFO and NYC markets, respectively. It appears that travelers in both markets reported similar experiences. Travelers mentioned Service Product (at about 26 %) and Physical Evidence (at about 25 %) most often, followed by Place (at

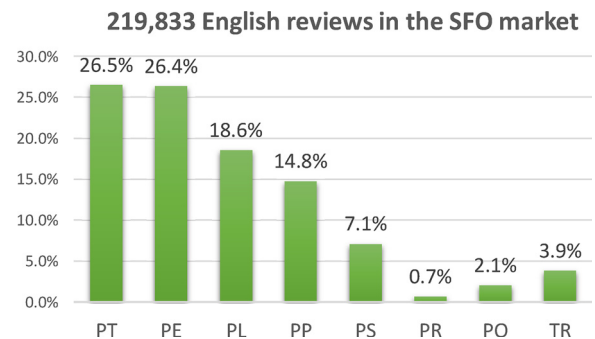


Fig. 2. The element distributions of the Airbnb reviews in the SFO market.

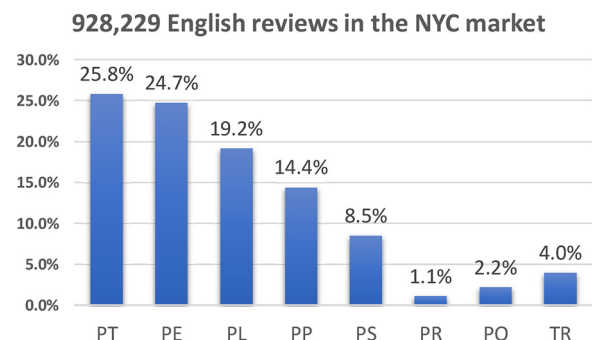


Fig. 3. The element distributions of the Airbnb reviews in the NYC market.

about 19 %) and Participant (at about 15 %). Service Process was the element that mentioned less than 10 %, with Promotion (at about 2%) and Price (at about 1%) listed as the least talked-about element.

Referring to annotation schema in Table 1, we concluded that travelers’ experience on Airbnb stays mainly reflect on their overall impressions or feelings about the intangible aspect of lodging product (Service Product) as well as Physical Evidence, such as the upkeep of the facility and the amenities provided in a listing. Travelers also pay attention to the accessibility of the location (Place) and hospitable hosts (Participant). Service Process is an area where travelers often mentioned the procedures for check-in or how the hosts handle customer service issues. For example, travelers stated: “Communication was about as prompt as it could be.” “We messaged [host] asking if we could borrow a pot for the kitchen and a minute later we heard a quiet clunk outside our door with a subsequent message.” “Problem solved!” “Their home guidelines and help to simply life.”

Interestingly, even though Promotion and Price are two critical elements in sales and marketing, travelers seldom mentioned these two Ps in their reviews in both markets. It is plausible that most descriptions and the pictures attached to the Airbnb listings genuinely reflect on what the travelers saw or experienced in person. As a result, they might not find the needs to make such comments as “the listing is pretty close to the images posted” unless they found themselves in a situation like “this place is even more adorable and quaint than the pictures show, if that’s possible :)” Meanwhile, even though price is a critical indicator of

a lodging product's service quality (Xie and Kwok, 2017), it is also a variable that the hosts can easily manipulate (Gibbs et al., 2018; Kwok and Xie, 2019). When travelers browse the available listings in a market, they have already seen the price and the fee structure of a listing. Before they request for an Airbnb stay, they have to agree on the price that they would be charged because the listed price is non-negotiable on Airbnb.com. Because travelers have already accepted and acknowledged the price before they travel, they might not see the need to overelaborate such a 7 P element anymore in their reviews. It is not surprising to see only a small number of travelers mentioned price in their comments. In the cases where the Price element was mentioned, travelers usually compared their Airbnb stay against a hotel stay, including "For the same price you would get a small hotel room downtown trapped in the business and tourist areas, which makes for less fun." "This was our first air bnb experience and we will now always use this site instead of booking expensive small hotel rooms." "Much less money than a hotel, and many more memories." Other travelers simply mentioned: "Well worth the price." or "Not worth the price."

4.2. The marketing mix of the listings managed by multi-unit hosts vs. those managed by single-unit hosts

To answer RQ2, we first divided the online reviews from both the SFO and NYC market according to the listings managed by multi-unit hosts and those managed by single-unit hosts. Because some reviews were not clearly labeled if they were written for the listings managed either type of hosts, possibly due to some programming issues when InsideAirbnb compiled the data for public access, we then removed those reviews without attaching to a listing managed by a multi-unit or a single-unit host from our dataset before performing additional analysis. Eventually, we retained a clearly labeled dataset with 166,191 reviews of the 3,108 listings that were managed either by multi-unit or single-unit hosts in the SFO market, plus another clearly labeled dataset with 813,821 reviews of the 28,661 listings that were managed either by multi-unit or single-unit hosts in the NYC market.

By comparing the element distributions shown in Figs. 4 and 5, it does not appear that the travelers staying in the listings managed by multi-unit hosts would share more (or fewer) comments on any particular marketing mix elements. Nevertheless, the element distributions, as shown in Figs. 4 and 5, demonstrate a similar pattern to what has been reported in Figs. 2 and 3.

Such findings may seem to contradict our discussion with relevant literature in Section 2.3, which suggests travelers might share different experiences for their stays with the multi-unit hosts as compared to the stays with the single-unit hosts. We then randomly selected 20 comments from each of the 7 Ps marketing mix elements for the listings managed by the multi-unit host plus 20 comments from each of the 7 Ps marketing mix elements for the listings managed by the single-unit hosts. Through a qualitative content analysis by the research team, we

found no observable differences in most elements except for the Participant element, where only 35 % of the travelers who stayed in a listing managed by a multi-unit host mentioned the host's name, but 55 % of the travelers who stayed in a listing managed by a single-unit host stated the host's name. It is possible that single-unit hosts have more face-to-face interactions with the travelers because they devote 100 % of their time and attention to one guest or one group of travelers at a given time, but multi-unit hosts must distribute their time and attention to multiple parties of travelers since they were managing more than one unit simultaneously. Table 3 provides some examples under the Participant element between the reviews written for the listings managed by single-unit hosts and the ones written for the listings managed by multi-unit hosts.

4.3. The marketing mix of the listings managed by the superhosts vs. those managed by the ordinary hosts

To answer RQ3, we regrouped the online reviews from both the SFO and NYC markets into two groups: reviews for the listings managed by the superhosts and those for the listings managed by the ordinary hosts. After removing the reviews with no clear labels, we retained a dataset with 166,489 reviews of the 3,128 listings in the SFO market and another dataset with 814,634 reviews of the 28,693 listings in the NYC market for analysis. On average, listings managed by superhosts received about twice as many reviews as the listings managed by the ordinary hosts (54 reviews per superhost listing as compared to 23 reviews per ordinary-host listing), indicating that listings managed by superhosts are more popular among travelers.

Fig. 6 and 7 reveal the "7 Ps + TR" element distributions between the superhosts and ordinary hosts in the SFO and NYC markets respectively. Both markets showed very similar patterns. When travelers described their experiences with an Airbnb listing that was managed by a superhost, they posted more comments regarding how much they enjoyed their Airbnb experience (Service Product), and at the same time, they thought highly about the Participant (the hosts). After travelers stayed with the ordinary hosts, they shared their experience using more words that describe the aspects of the listings' Physical Evidence and Place (location). Our results only show minimal or even no gaps between these two groups under the Service Process, Price, Promotion, and Traveler elements.

In the environments where hosts and travelers rate their experience against the other party, the social pressure they received would push them to write more positive comments about each other (Botsman, 2017). Under such social pressure, people might not want to explicitly discuss their negative or not-as-pleasant experience. Instead, they would probably focus on sharing the positive experience in their reviews. If that is the case, what they do not state in their reviews could also serve as a good indicator of the service received. When travelers wrote less about their overall experience with their stays (Service

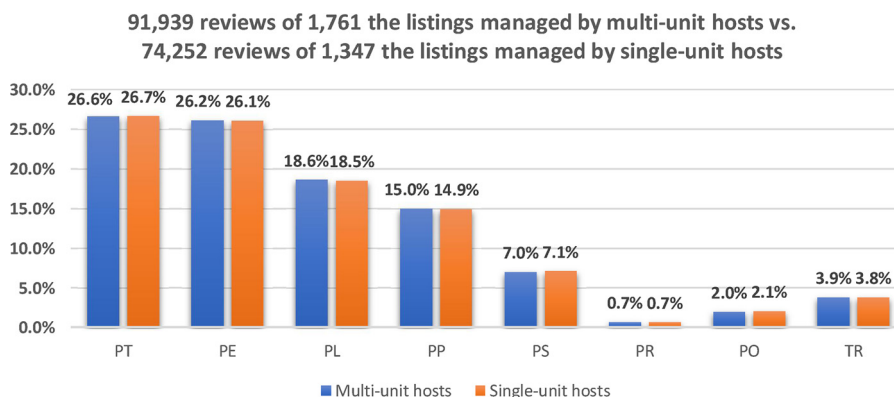


Fig. 4. Airbnb reviews' element distributions for the listings managed by multi-unit and the listings managed by single-unit hosts in the SFO market.

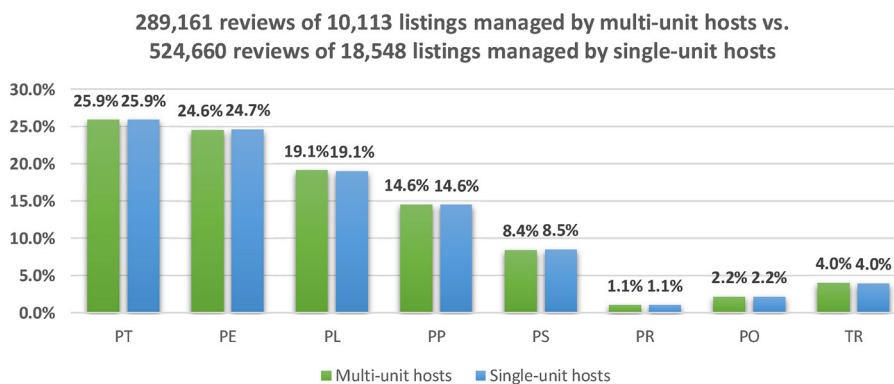


Fig. 5. Airbnb reviews' element distributions for the listings managed by multi-unit and the listings managed by single-unit hosts in the NYC market.

Table 3
Examples under the Participant element.

Single-unit host	Multi-unit host
<i>[host name 1] and [host name 2] were amazing!</i>	<i>Owners are great to communicate with and go out of their way to make your stay comfortable.</i>
<i>[host name 1] and [host name 2] were very friendly and made sure I had everything I needed.</i>	<i>They helped us with everything we needed and were able to give us some good suggestions about where to go and what to see in the Bay Area.</i>
<i>[host name] was extremely gracious and welcoming and was prompt with all communications.</i>	<i>My wife and I didn't have a chance to meet the owners, but we did see their assistant [the name of the assistant to the multi-unit host].</i>

Product) or about the hospitable hosts (Participant) for their stays with the ordinary hosts but focused more on the Physical Evidence and the location (Place) of the listings, such element distributions might indicate that the superhosts indeed outperformed the ordinary hosts by offering better overall service and making travelers feel more welcome.

We went beyond the distributions to further examine the review content under the superhost group to check how the content might differ from the reviews written for the ordinary hosts or not. As illustrated in the specific examples from the reviews in Table 4, we found the comments were overwhelmingly positive or showed stronger positivity across all the 7 Ps categories for the superhosts. We also observed more lukewarm or sometimes even negative comments for the ordinary hosts. Such evidence further confirms the notion that superhosts indeed provided better services than the ordinary hosts.

5. Implications and Conclusion

This study reports a big-data, supervised machine learning analysis that was built upon the 7 Ps marketing mix framework in service marketing. Through our analysis of 219,833 English Airbnb reviews in SFO as well as 928,229 English Airbnb reviews in NYC, we gained a better understanding of the marketing mix of Airbnb listings — the revolutionary newcomers to the lodging industry. The insights revealed

in this study provide useful BI for related businesses as they make managerial decisions, such as in differentiation and positioning. This study's theoretical and practical implications warrant additional discussion.

5.1. Theoretical implications

First and for most, while there were critiques and debates of whether the marketing mix framework was still relevant to the marketing practices in the 21st Century (e.g., Constantinides, 2006; Möller, 2006), our analysis provides a great example of how the 7 Ps Model can remain to be a versatile framework even in the era where businesses heavily rely on the immense big data available on the internet for BI. Not only the 7 Ps Model enabled us to break down the marketing mix elements of travelers' experience on Airbnb, but it also allowed us to systematically compare and contrast travelers' Airbnb experience with different types of service products managed by various service providers (types of hosts). This study complements the relevant studies that also adopted the text-mining approach to analyze consumer reviews on Airbnb (e.g., Bridges and Vásquez, 2018; Cheng and Jin, 2019).

Second, we contribute to the literature in information technology and information management by validating a newly-introduced supervised machine learning technique in analyzing the online review

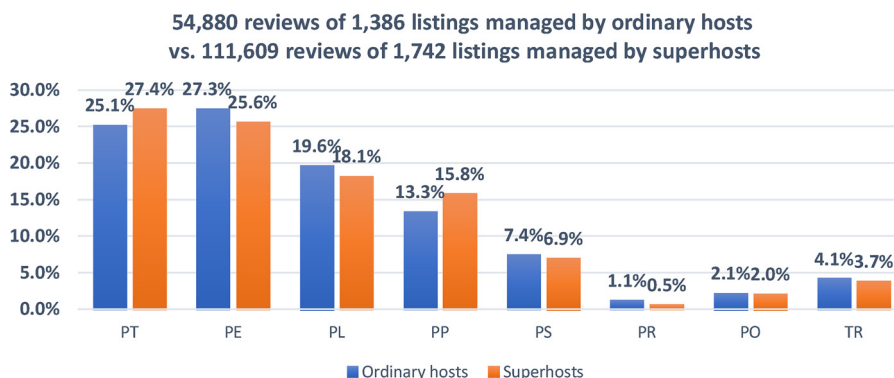


Fig. 6. Element distributions between the listings managed by superhosts and those managed by ordinary hosts in SFO.

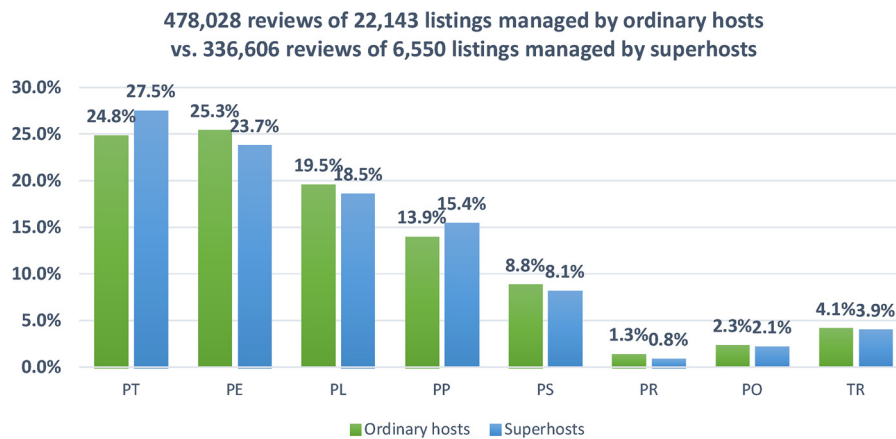


Fig. 7. Element distributions between the listings managed by the superhosts and those managed by ordinary hosts in NYC.

data that are accessible for almost anyone and every business. We echo other scholars' suggestions regarding online review studies that technology-enabled big data analytics will provide different and valuable insights from the conventional survey-based metrics or experiments (e.g., Kwok et al., 2017; Xiang et al., 2015). We are hoping our study will inspire more researchers in relevant disciplines to take advantage of the publicly-accessible data online in their journey of discovering new insights into a complex phenomenon. This study joins the relevant scholarly discussion (e.g., Fan et al., 2015; Zhao et al., 2019; Xiang et al., 2015) by demonstrating an additional example of using big data analytics for BI.

Lastly, while adding new knowledge to current literature might not be considered as a significant contribution to theory building, our analysis helps discover new insights about home-sharing services. Supplementary relevant practical implications can hence be advanced.

5.2. Practical Implications

The 4 Ps Model is particularly useful in helping businesses develop the optimal marketing mix for market segmentation, differentiation, and positioning (Möller, 2006). Because of this study's unique research setting, the webmasters and the hosts with an entrepreneurial mindset can benefit from the research findings. For example, the webmasters of home-sharing websites can adopt the algorithm introduced in this

report to assess the service quality of the potential superhosts in addition to the standards outlined in Airbnb's website. Our algorithm also provides a useful tool for home-sharing websites to conduct additional analyses, such as analyzing the 7 Ps marketing mix by different market segments, allowing them to develop differentiation strategies in market positioning.

Then, the hosts who are striving to be recognized as the superhosts are highly encouraged to review the results reported Section 4.3 and Table 4 carefully. With a better understanding of how superhosts differentiate their services from those offered by the ordinary hosts, hosts can mirror what the superhosts do in each of the 7 P elements and adjust their offerings accordingly. For example, because it may seem travelers appreciate superhosts' involvement in the service process by mentioning the "Participant" more often in their comments than what they would write for their stays with the ordinary hosts. It becomes critical for the ordinary hosts to find ways to interact with their guests a little bit more, such as doing a tour of the home-sharing facility by themselves as a means to welcome the guests.

Additionally, we highly encourage the hoteliers who have already entered to home-sharing market or want to develop effective strategies to combat the threat created by Airbnb and other home-sharing services to revisit the 7 Ps marketing mix of their own products. By comparing what hotels offer and what Airbnb listings offer in terms of the 7 Ps marketing mix, hoteliers will very likely be able to see a clearer picture

Table 4

Selective comments for the superhosts and the ordinary hosts.

Element	Superhosts	Ordinary Hosts
Service Product (PT)	<i>Airbnb community is lucky to have [superhost name] as a new host and would love to stay with her again anytime! I highly recommend [superhost name] and her home if you plan on visiting San Fran.</i>	<i>It's a good place to stay. Anyhow, would not recommend this to anyone... just dirty and dingy.</i>
Price (PR)	<i>Well worth a stay!</i>	<i>All in all, we think this space has potential, but would not recommend it as-is for the price. Overall it's ok but I thought the price was high for the quality of the apartment.</i>
Place (PL)	<i>The location (Richmond District) offers great eats and puts you in close proximity to some amazing SF sites. The neighbourhood is lovely as well, quiet with nice places to eat nearby and of course next to the amazing park.</i>	<i>The apartment is in a great location - I was there for the Outside Lands festival, and it was only few blocks away from Golden Gate Park.</i>
Promotion (PO) Partici-pants (PP)	<i>[super host name]'s house is better than looks on the photos. [super host name] is an amazing host! [super host name] is extremely hospitable and she's happy to share a glass of wine with at the end of the day and hear about your adventures.</i>	<i>The apartment was clean and accurate to description. [host name] is very nice. [host name], the host, was very friendly and helpful.</i>
Physical Evidence (PE)	<i>The entire house is beautiful and the bathroom and bed room are super clean, very nice and comfortable! There are many modern touches that blend in seamlessly with the original character of the house.</i>	<i>The studio is simple, but more than sufficient for a nice stay. While the cabinets and countertops in the kitchen area were nice and newer, the appliances were very, quite old and a little bit scary to use, especially the toaster oven and hot plate.</i>
Service Process (PS)	<i>He even arranged to have us check in early which was very much appreciated.</i>	<i>She replied our message quickly before we arrived and warm-hearted to solve our problems during our stay.</i>

of what they did well or what they did poorly. They will also likely be able to come up with specific marketing and operational strategies, allowing them to differentiate themselves by focusing on one or two aspects of the marketing mix elements.

Last but not least, we echo other researchers that the policymakers should pay close attention to the multi-unit hosts as they propose new policies to restrict the short-term residential rentals in a market (e.g., Horn and Merante, 2017; Kwok and Xie, 2019; Xie et al., 2020a-b). Multi-unit hosts have already gained many operational advantages over other hotels or hostels because they normally need not follow the legal and safety compliances as what the regular hotels do. Furthermore, current literature reported that multi-unit hosts could gain substantially higher revenues than the single-unit hosts (e.g., Magno et al., 2018; Wegmann and Jiao, 2017; Kwok and Xie, 2019). Our findings also suggest that travelers enjoy a very similar experience from the listings managed by multi-unit and single-unit hosts. It may seem that the multi-unit hosts as micro-entrepreneurs, instead of those mom-and-pop-like single-unit hosts, could be a more significant threat to the traditional hotels or hostels and thus deserve more attention from the policymakers.

6. Limitations and Future Research

This study is not without limitations. First, even though we applied our tested algorithms to analyze the online reviews from two entire major markets for short-term residential rental business in the U.S., our results may not be easily generalized to other global markets, primarily when our analysis only focused on the online reviews in English. Future studies may consider tuning our algorithm in different languages, allowing for additional textual analyses on the online reviews written in other languages than English. Second, it is important to note that even though the machine learning approaches usually allow researchers to automatically analyze the immense amount of the big data available on the internet, no automatic analyses are free of errors. Likewise, our method is not perfect even though we took the best measures we could do to avoid errors in our analysis with detailed step-by-step descriptions. As we adjusted the definition of the 7 Ps elements specifically for the Airbnb listings to avoid bad operationalization of our analysis, extra cautions should be given when interpreting the results that are not error-free. Hence, new methods or investigations under a different theory are highly encouraged in future research because studies that utilize multiple theoretical frameworks with different methodologies are very likely to bring in new insights to our understanding of a complex phenomenon (Kwok, 2012). Third, our assessment only focuses on comparing the marketing mix of the listings managed by multi-unit vs. single-unit hosts and those managed by superhosts vs. ordinary hosts. Using the 7 Ps model, Airbnb experiences between first-time users and repeat customers can also be compared, of which the results can join the scholarly discussion about travelers' repeat purchase behaviors on home-sharing websites (e.g., Xie et al., 2019; Xie et al., 2020a). Lastly, this study reports an analysis based on the data collected before the COVID-19 pandemic, when people did not show much concern about the coronavirus or the hygiene standards of a lodging facility, consumers' emphasis of the 7 P elements could be shifted. We highly recommend future studies to examine how consumers' comments about the 7 P elements have changed in the post – COVID-19 era.

In conclusion, we adopted a new big-data analytic approach in analyzing the online review data on Airbnb. Our findings, even though not without limitations, add to relevant literature and benefit the webmasters and hosts on home-sharing websites, as well as the hoteliers who are finding effective ways to combat the threat created by the still-growing home-sharing sector. We also encourage future studies in service research to take advantage of what the publicly-accessible data may offer.

Declaration of Competing Interest

The authors declare no conflicts of interest with respect to the authorship and/or publication of this article.

Acknowledgement

This study was supported with an internal fund made available by the Collins College of Hospitality Management at Cal Poly Pomona.

References

- Abrate, G., Viglia, G., 2019. Personal or product reputation? Optimizing revenues in the sharing economy. *J. Travel. Res.* 58 (1), 136–148.
- Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., Weaven, S., 2019. Market segmentation and travel choice prediction in spa hotels through tripadvisor's online reviews. *Int. J. Hosp. Manag.* 80, 52–77.
- Airbnb. Superhost: Recognizing the best in hospitality. (n.d.) Retrieved on April 10, 2019 via <https://www.airbnb.com/superhost>.
- Blal, I., Singal, M., Templin, J., 2018. Airbnb's effect on hotel sales growth. *Int. J. Hosp. Manag.* 73, 85–92.
- Booms, B.H., Bitner, M.J., 1981. Marketing strategies and organization structures for service firms. In: Donnelly, J.H., George, W.R. (Eds.), *Marketing of Services*. American Marketing Association, Chicago, IL, pp. 47–51.
- Borden, N.H., 1964. The concept of the marketing mix. *J. Advert. Res.* 4 (2), 2–7.
- Botsman, R., 2017. *Who Can You Trust? How Technology Brought Us Together – and Why It Could Drive Us Apart*. Penguin, UK.
- Bridges, J., Vázquez, C., 2018. If nearly all Airbnb reviews are positive, does that make them meaningless? *Curr. Issues Tour.* 21 (18), 2057–2075.
- Cheng, M., Jin, X., 2019. What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* 76, 58–70.
- Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R., 2011. Signaling theory: a review and assessment. *J. Manage.* 37 (1), 39–67.
- Constantinides, E., 2006. The marketing mix revisited: towards the 21st century marketing. *J. Mark. Manag.* 22 (3-4), 407–438.
- Cui, R., Li, J., Zhang, D.J., 2017. Discrimination with Incomplete Information in the Sharing Economy: Field Evidence from Airbnb. Available at <https://ssrn.com/abstract=2882982>.
- El-Bawab, N., 2019. Marriott plans to launch home-rental market platform that would compete with Airbnb, report says. CNBC News Retrieved on March 1, 2020 via <https://www.cnbc.com/2019/04/29/marriott-to-launch-home-rental-platform-to-compete-with-airbnb-report.html>.
- Fader, P.S., Winer, R.S., 2012. Introduction to the special issue on the emergence and impact of user-generated content. *Mark. Sci.* 31 (3), 369–371.
- Fan, S., Lau, R.Y.K., Zhao, J.L., 2015. Demystifying big data analytics for business intelligence through the lens of marketing mix. *Big Data Res.* 2, 28–32.
- Feng, S., Zhao, C., Fu, P., 2020. A deep neural network based hierarchical multi-label classification method. *Rev. Sci. Instrum.* 91 (2), 024103.
- Fong, S.W., Skillicorn, D.B., Roussinov, D., 2006. Detecting word substitution in adversarial communication. 6th SIAM Conference on Data Mining. Bethesda, Maryland.
- Gibbs, C., Guttentag, D., Gretzel, U., Yao, L., Morton, J., 2018. Use of dynamic pricing strategies by Airbnb hosts. *Int. J. Contemp. Hosp. Manage.* 30 (1), 2–20.
- Goi, C.L., 2009. A review of marketing mix: 4Ps or more? *Int. J. Mark. Stud.* 1 (1), 2.
- Google, 2019. Google Cloud Translation API: Detecting Languages. Retrieved from <https://cloud.google.com/translate/docs/detecting-language> on March, 27th, 2019.
- Guess Language, 2019. Project Description. Retrieved from <https://pypi.org/project/guess-language-spirit/> on March 27th, 2019.
- Guttentag, D., 2015. Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Curr. Issues Tour.* 18 (12), 1192–1217.
- Hanssens, D.M., Pauwels, K.H., Srinivasan, S., Vanhuele, M., Yildirim, G., 2014. Consumer attitude metrics for guiding marketing mix decisions. *Mark. Sci.* 33 (4), 534–550.
- Harrington, R.J., Ottenbacher, M.C., Fauser, S., 2017. QSR brand value: marketing mix dimensions among McDonald's, KFC, Burger King, subway and Starbucks. *Int. J. Contemp. Hosp. Manage.* 29 (1), 551–570.
- Heo, C.Y., Blengini, I., 2019. A macroeconomic perspective on Airbnb's global presence. *Int. J. Hosp. Manag.* 78, 47–49.
- Horn, K., Merante, M., 2017. Is home sharing driving up rents? Evidence from Airbnb in Boston. *J. Hous. Econ.* 38, 14–24.
- Jiang, Y., Balaji, M.S., Jha, S., 2019. Together we tango: value facilitation and customer participation in Airbnb. *Int. J. Hosp. Manag.* 82, 169–180.
- Joachims, T., 1998. Text categorization with support vector machines: learning with many relevant features. In: *European Conference on Machine Learning*. Springer, Berlin, Heidelberg, pp. 137–142.
- Ju, Y., Back, K.J., Choi, Y., Lee, J.S., 2019. Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *Int. J. Hosp. Manag.* 77, 342–352.
- Karlsson, L., Kemperman, A., Dolnicar, S., 2017. May I sleep in your bed? Getting permission to book. *Ann. Tour. Res.* 62, 1–12.
- Kwok, L., 2012. Exploratory-triangulation design in mixed methods studies: a case of examining graduating seniors who meet hospitality recruiters' selection criteria. *Tour. Hosp. Res.* 12 (3), 125–138.
- Kwok, L., Xie, K.L., 2016. Factors contributing to the helpfulness of online hotel reviews:

- Does manager response play a role? *Int. J. Contemp. Hosp. Manage.* 28 (10), 2156–2177.
- Kwok, L., Xie, K.L., 2018. Buyer-seller similarity: does it lead to a successful transaction of peer-to-peer home-sharing services? *Int. J. Contemp. Hosp. Manage.* 30 (9), 2925–2944.
- Kwok, L., Xie, K.L., 2019. Pricing strategies on Airbnb: are multi-unit hosts revenue pros? *Int. J. Hosp. Manag.* 82, 252–259.
- Kwok, L., Xie, K.L., Richards, T., 2017. Thematic framework of online review research: a systematic analysis of contemporary literature on seven major hospitality and tourism journals. *Int. J. Contemp. Hosp. Manage.* 29 (1), 307–354.
- Loo, P.T., Leung, R., 2018. A service failure framework of hotels in Taiwan: adaptation of 7Ps marketing mix elements. *J. Vacat. Mark.* 24 (1), 79–100.
- Lui, M., Baldwin, T., 2012. *Langid.py: an off-the-shelf language identification tool*. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012). Demo Session, Jeju, Republic of Korea. Available from: www.aclweb.org/anthology/P12-3005.
- Luo, Y., Tang, R.L., 2019. Understanding hidden dimensions in textual reviews on Airbnb: an application of modified latent aspect rating analysis (LARA). *Int. J. Hosp. Manag.* 80, 144–154.
- Magno, F., Cassia, F., Ugolini, M.M., 2018. Accommodation prices on Airbnb: effects of host experience and market demand. *TQM J.* 30 (5), 608–620.
- Mariani, M., Baggio, R., Fuchs, M., Höpken, W., 2018. Business intelligence and big data in hospitality and tourism: a systematic literature review. *Int. J. Contemp. Hosp. Manag.* 30 (12), 3514–3554.
- Mauri, A.G., Minazzi, R., Nieto-García, M., Viglia, G., 2018. Humanize your business. The role of personal reputation in the sharing economy. *Int. J. Hosp. Manag.* 73, 36–43.
- McCarthy, E.J., 1964. *Basic Marketing*. Richard D. Irwin, IL.
- Möller, K., 2006. The marketing mix revisited: towards the 21st century marketing by E. Constantinides. *J. Mark. Manag.* 22 (3-4), 439–450.
- Netzer, O., Feldman, R., Goldenberg, J., Fresko, M., 2012. Mine your own business: market-structure surveillance through text mining. *Mark. Sci.* 31 (3), 521–543.
- Padma, P., Ahn, J., 2020. Guest satisfaction & dissatisfaction in luxury hotels: an application of big data. *Int. J. Hosp. Manag.* 84, 102318.
- Pustejovsky, J., Stubbs, A., 2012. *Natural Language Annotation for Machine Learning: a Guide to Corpus-building for Applications*. O'Reilly Media, Inc.
- Rafiq, M., Ahmed, P.K., 1995. Using the 7Ps as a generic marketing mix: an exploratory survey of UK and European marketing academics. *Mark. Intell. Plan.* 13 (9), 4–15.
- Talón-Ballester, P., González-Serrano, L., Soguero-Ruiz, C., Muñoz-Romero, S., Rojo-Álvarez, J.L., 2018. Using big data from customer relationship management information systems to determine the client profile in the hotel sector. *Tour. Manag.* 68, 187–197.
- Ting, D., 2017. Airbnb ramps up push to get more hosts to choose instant booking. Skift Retrieved on April 10, 2019 via: <https://skift.com/2017/06/14/airbnb-ramps-up-push-to-get-more-hosts-to-choose-instant-booking/>.
- Wang, J.H., Wang, H.Y., Chen, Y.L., Liu, C.M., 2015. A constructive algorithm for un-supervised learning with incremental neural network. *J. Appl. Res. Technol.* 13 (2), 188–196.
- Wegmann, J., Jiao, J., 2017. Taming Airbnb; toward guiding principles for local regulation of urban vacation rentals based on empirical results from five US cities. *Land Use Policy* 69, 494–501.
- Wiles, A., Crawford, A., 2017. Network hospitality in the share economy: understanding guest experiences and the impact of sharing on lodging. *Int. J. Contemp. Hosp. Manag.* 29 (9), 2444–2463.
- Xiang, Z., Schwartz, Z., Gerdes Jr, J.H., Uysal, M., 2015. What can big data and text analytics tell us about hotel guest experience and satisfaction? *Int. J. Hosp. Manag.* 44, 120–130.
- Xie, K.L., Kwok, L., 2017. The effects of Airbnb's price positioning on hotel performance. *Int. J. Hosp. Manag.* 67, 174–184.
- Xie, K.L., Kwok, L., Wu, J., 2019. Are consumers loyal to home-sharing services? Impacts of host attributes and frequency of past stays. *Int. J. Contemp. Hosp. Manag.* 31 (3), 1066–1085.
- Xie, K.L., Kwok, L., Chen, C., Wu, J., 2020a. To share or to access? Travelers' choice on the types of accommodation-sharing services. *J. Hosp. Tour. Manag.* 42, 77–87.
- Xie, K.L., Kwok, L., Heo, C.Y., 2020b. Are neighbors friends or foes? Assessing Airbnb listings' agglomeration effect in New York City. *Cornell Hosp. Q.* 61 (2), 128–141.
- Yang, S.B., Lee, K., Lee, H., Koo, C., 2019. In Airbnb we trust: understanding consumers' trust-attachment building mechanisms in the sharing economy. *Int. J. Hosp. Manag.* 83, 198–209.
- Yin, H., Yao, T., 2008. Topic identification based on Chinese domain-specific subjective sentence. *Autonomous Systems-Self-Organization, Management, and Control*. Springer, Dordrecht, pp. 177–183.
- Zhao, Y., Xu, X., Wang, M., 2019. Predicting overall customer satisfaction: big data evidence from hotel online textual reviews. *Int. J. Hosp. Manag.* 76, 111–121.
- Linchi Kwok, Ph.D.**, is an associate professor in the Collins College of Hospitality Management at California State Polytechnic University Pomona (Cal Poly Pomona). He received an M.S. and a Ph.D. degree in Hospitality Administration at Texas Tech University, as well as an MBA at Syracuse University. His research interests include information technology and service operations.
- Yingying Tang, M.S.**, recently graduated from the Business Analytics Program in the Whitman School of Management at Syracuse University, where she received a Master of Science degree. Her interest is in business intelligence and data mining.
- Bei Yu, Ph.D.**, is an Associate Professor at the School of Information Studies, Syracuse University. Her research expertise is in applied natural language processing and computational social science.