

Article

The Determinants of Helpful Hotel Reviews: A Social Influence Perspective

Yukyung Son ¹, Kyungmo Kang ², Ilyoung Choi ^{2,*} and Jaekyeong Kim ^{2,3}

¹ Department of Business Administration, Graduate School, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02453, Republic of Korea

² Department of Big Data Analytics, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02453, Republic of Korea

³ School of Management, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02453, Republic of Korea

* Correspondence: choice102@khu.ac.kr; Tel.: +82-2-961-9355

Abstract: Online hotel reviews regarding specific experiences and the sensibility of hotel reviewers are important sources of information for consumers who want to book hotels in the future. However, since thousands of reviews are written for one hotel, it is practically impossible for customers to read all the reviews. To alleviate this problem, OTAs (online travel agencies) provide a helpful vote of reviews, helping customers to quickly find helpful reviews without much effort. Therefore, in this study, a ZINB (zero-inflated negative binomial regression) was performed to investigate factors that influence the helpfulness of hotel reviews using the social influence theory. As a result of the analysis, it was found that location, service, and value of the hotel, alongside normative influencing factors and the length of the review (including informational influencing factors), affect the helpfulness of the review regardless of the city. The results of this study are expected to help hotel managers to take remedial action on negative reviews and strengthen the promotion of positive reviews, potentially helping to increase customer satisfaction.



Citation: Son, Y.; Kang, K.; Choi, I.; Kim, J. The Determinants of Helpful Hotel Reviews: A Social Influence Perspective. *Sustainability* **2022**, *14*, 14881. <https://doi.org/10.3390/su142214881>

Academic Editor: Jun (Justin) Li

Received: 14 October 2022

Accepted: 9 November 2022

Published: 10 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: social influence theory; online review; helpfulness; zero-inflated negative binomial regression

1. Introduction

As the hotel industry is connected to information and communication technology, OTAs (online travel agencies), such as TripAdvisor.com, Expedia.com, and Booking.com, are being activated. Approximately 57% of hotel reservations are made online; therefore, OTAs are essential for the hotel industry. OTAs provide online reviews as well as hotel information, such as rooms, meals, banquet facilities, prices, and hotel ratings in various languages.

In addition to the overall rating, the online hotel review consists of a textual form indicating the rating of cleanliness, location, rooms, service, sleep quality, value, and subjective (positive, neutral, and negative) feedback on the hotel, indicating the customer's opinion of the hotel [1]. Online reviews of these hotels are directly exposed to the public so that customers who want to use the hotel can easily access them. Therefore, online reviews affect hotel reservations because customers who want to use a hotel can easily compare them with other hotels through online reviews [2–5]. In other words, online reviews of hotels affect business performance [6,7] and reputation [8], including hotel sales.

However, since thousands of reviews are written for one hotel, it is practically impossible for customers to read all reviews. OTAs provide “helpful votes” of reviews to alleviate this problem, which helps customers to quickly see hotel evaluations without much effort. Therefore, the helpfulness of reviews plays a vital role in enabling customers who want to use the hotel to filter reviews efficiently [9].

Accordingly, many studies have analyzed the factors that influence the helpfulness of reviews on websites such as Amazon.com [10–12], Yelp.com [13,14], and TripAdvisor.com [15,16], and found that the helpfulness of reviews is affected by the attributes and behaviors of other reviewers, such as rating and number of friends. However, although many studies have examined the factors that influence the helpfulness of reviews, no studies have determined whether multi-attribute ratings (e.g., ratings for cleanliness, location, rooms, service, sleep quality, and value in hotel reviews) affect the helpfulness of reviews.

Therefore, in this study, the social influence theory was used to find factors influencing the helpfulness of hotel reviews. To this end, we reviewed data from hotels located in New York, Seoul, and Shanghai from TripAdvisor. The reviewed data considers of cleanliness, location, rooms, service, sleep quality, value ratings, review text, helpful votes, and hotel grade. The moderating effect of hotel grade (luxury hotel vs. budget hotel) was investigated by dividing hotels into luxury hotels and budget hotels using hotel ratings.

The results of this study provide strong empirical evidence for evaluating hotel reviews. We also found that there were differences in the multiple attributes affecting review helpfulness according to the grade (i.e., star rating) and geographic location of the hotel. Therefore, the results of this study can make theoretical and practical contributions to the hotel industry.

2. Research Background

2.1. Social Influence Theory

People change their thoughts, feelings, and behaviors by being influenced by people around them and their beliefs and attitudes. In other words, changes in human thinking, emotion, and behavior occur under the influence of others. Social influence theory explains the interpersonal process that changes these group members' thoughts, feelings, and behavior. Humans are largely affected by normative influence and informational influence of others to change behavior [17]. Normative influence means changing one's behavior according to one's specific expectations, attitudes, and behaviors to be liked or recognized by another person. Since humans are social animals, they desire to belong to influential and vital groups rather than living alone. Customers who respond to normative influences tend to meet other people's expectations through product or service purchases.

On the other hand, informational influence refers to imitating behavior by referring to information obtained by others. In other words, in situations where one cannot trust one's judgment, people act according to other people's opinions to make better decisions. Therefore, customers tend to collect information from others and make decisions.

Online reviews affect people who read reviews. Accordingly, many researchers have studied the social influence of online reviews. The authors of [18] investigated the moderating effects of experience goods or search goods to explore factors determining a review's helpfulness. As a result, they determined that reviewer reputation, identity disclosure, and review depth positively affect the helpfulness of online reviews. Furthermore, they discovered that reviewer identity disclosure has a more significant effect on review helpfulness of experience goods. On the other hand, it was found that the reviewer's reputation, the review's degree, and the review's depth affect the review helpfulness of search goods.

The authors of [19] investigated the effect of activity level and review helpfulness on customer influence. The study found that activity level had a greater influence on customers with high strength of transactional connectivity than customers with low strength, and the impact of customer review helpfulness on influence was stronger for people with high similarity than those with low similarity. The authors of [20] investigated the effect of helpfulness votes on the customer's perceived helpfulness for the review and whether social influence moderates the relationship between content expression factors and perceived helpfulness. The analysis revealed that reviews with clear positive (or negative) helpfulness votes are more (or less) helpful. In contrast, the helpfulness votes of others reduce the positive effects of structures and the negative effects of spelling errors. The authors of [14] investigated the effect of the order of reviews on the review's helpfulness from a social

influence perspective and found that the order of reviews not only negatively affects the helpfulness of reviews but also weakens the order effect for negative reviews.

Existing studies used review attributes and reviewer attributes to identify determinants of review helpfulness. In particular, only the overall rating was used as a factor for normative influence. However, customers evaluate hotels in multiple dimensions, such as price and location, and other customers use this information to make purchase decisions. Therefore, this study tried to identify the determinants of review helpfulness using the multiple attribute as factors for the normative influence.

2.2. Expectation Disconfirmation Theory

The expectation disconfirmation theory (EDT) is a widely used theory to describe the process of customer satisfaction formation [21]. EDT forms satisfaction or dissatisfaction with the product or service through the following steps [22]. The first step is for the customer to create pre-purchase expectations for a particular product or service. The second step is to evaluate the quality after purchasing a product or service. The third step is to compare the perceived quality with pre-purchase expectations. If the perceived quality is higher than or equal to expectations, the customer is satisfied, but if the perceived quality is lower than expected, the customer is dissatisfied. In the final step, satisfied customers form a willingness to repurchase and spread positive rumors, but dissatisfied customers avoid repurchasing and spread negative rumors.

Therefore, customers have different expectations according to the hotel grade. Customers generally expect high service quality for luxury hotels but relatively low service quality for budget hotels. If the hotel's service quality is higher than expected, the customer will be satisfied and write a positive review. On the other hand, if the hotel's service quality is lower than expected, the customer will be dissatisfied and write negative reviews. These reviews influence customers' decisions to use the hotel in the future.

3. Research Model and Hypothesis Development

In this study, we examine the causal relationship between factors for normative influence (cleanliness, location, rooms, service, sleep quality, and value), factors for informational influence (number of words and sentiment), and review's helpfulness according to hotel grade.

From this point of view, cleanliness, location, room, service, sleep quality, and value were regarded as normative influencing factors in this study. The number of review words and review sentiment were regarded as informational influencing factors. In addition, this study aims to explain the factors that influence decision-making according to hotel grades using TripAdvisor.com data. Therefore, the model of this study is shown in Figure 1.

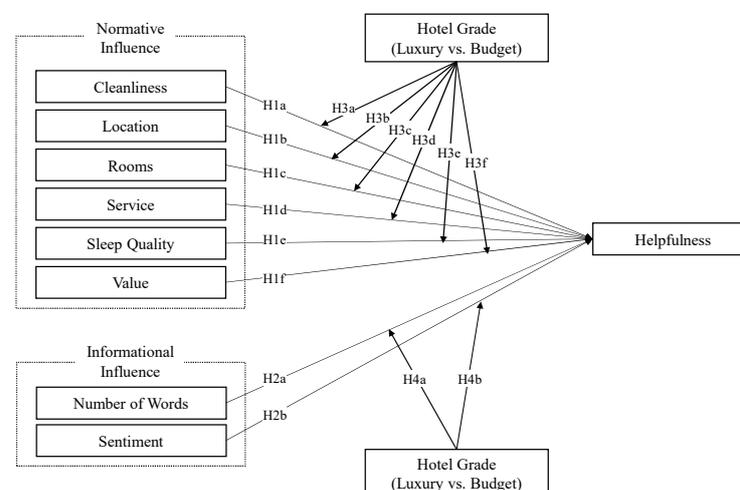


Figure 1. Research Model.

3.1. Normative Influencing Factors: Cleanliness, Location, Rooms, Service, Sleep Quality, Value

Normative influence is other reviewers' tendency to conform to previous reviewers' expectations [14,17,23–25]. Therefore, the normative influence refers to the opinions of other reviewers, such as review ratings [2]. The review rating represents the overall satisfaction with a product or service marked with stars (one to five points in numbers) [26]. In other words, reviews with fewer stars mean that reviewers are not satisfied with the product or service, and reviews with more stars mean that they are satisfied with the product or service. Therefore, other customers determine whether a product or service is helpful through the ratings from the previous reviewer [9,10,27].

Many studies have investigated whether review ratings affect helpfulness, and the relationship between review ratings and helpfulness shows largely positive, negative, U-shaped, and inverted U-shaped relationships. For example, [11,28,29] argued that review ratings had a positive effect on helpfulness. However, [30,31] argued that review ratings had a negative effect on helpfulness. On the other hand, [16] argued that the relationship between review rating and helpfulness was a U-shape, and [32] argued that it was an inverted U-shape relationship.

When customers leave a review of the hotel, they evaluate not only the overall rating of the hotel but also six attributes such as cleanliness, location, room, service, sleep quality, and value. The ratings for six attributes contain more specific information about the hotel. For example, it can be inferred that a hotel with a high rating for its location has a good view even if customers do not read the reviews. In addition, customers have various preferences for the six attributes of a hotel. For example, some customers will value the location of a hotel, while others will value the cleanliness of the hotel. Thus, ratings on the six attributes help determine whether a hotel review is as helpful as the overall rating. According to the above discussion, this study proposes the following hypotheses.

H1a. *The hotel's cleanliness will significantly affect the review's helpfulness.*

H1b. *The hotel's location will significantly affect the review's helpfulness.*

H1c. *The hotel's rooms will significantly affect the review's helpfulness.*

H1d. *The hotel's service will significantly affect the review's helpfulness.*

H1e. *The hotel's sleep quality will significantly affect the review's helpfulness.*

H1f. *The hotel's value will significantly affect the review's helpfulness.*

3.2. Informational Influencing Factors: Review Length (Number of Words), Sentiment

Informational influence refers to other reviewers' tendency to conform to previous reviewers' opinions based on information, arguments, facts, etc., obtained from previous reviews [14,17,33]. Customers refer to previous reviews of the product or service as it is assumed that the previous reviewer had a better understanding of the product or service [14]. In general, informational influences include argument quality, recommendation framing, etc. [26].

An argument is a piece of information in a review [34], and the quality of the argument is the strength of the message in the review [26,35]. The quality of the argument was measured using the review length, as the number of messages contained in the review helped determine whether the product or service was helpful [10,31,35–39].

Many studies have investigated whether review length affects helpfulness, and the relationship between review length and helpfulness shows a positive or negative relationship. For example, [32] argued that helpfulness decreases as the number of words included in the review of experience and search goods increases. In addition, [13,14,31] argued that increasing the length of the review increases the helpfulness on Yelp.com. The authors of [11] argued that the longer the review length for cell phone, printer, camera, music player, music cd, and video game, the more helpful it is. That is, the review length is related to the richness of hotel information [13]. As longer reviews are more likely to describe the

customer's hotel experience in detail, a longer review helps customers to know more about the hotel. Therefore, in this study, the hypothesis was established as follows.

H2a. *Review length will significantly affect the review's helpfulness.*

Recommendation framing is defined as positive, neutral, and negative emotions of the message implied in the review [26]. Methods of grasping the emotions of the review are largely LIWC (linguistic inquiry and word count)-based emotion analysis and machine learning-based emotion analysis methods. First, LIWC-based emotion analysis is a method of calculating the ratio of words representing positive emotion and negative emotion within a review and is used in online reviews such as MOOC review [40], hotel booking review [41], restaurant reviews [42] and Amazon product review [43]. Second, machine learning-based emotion analysis is a method of classifying emotions between -1 (negative emotion) and 1 (positive emotion) using an SVM (support vector machine), naive Bayes classifiers, and deep learning, and these methods are mainly performed in algorithmic research.

Customers describe their experiences of the hotel through reviews. In general, customers with good experiences write positive reviews, while customers with bad experiences write negative reviews. As many previous studies have suggested, the review's sentiment significantly affects its helpfulness [15,44,45]. In other words, because customers have an indirect experience of hotel services through reviews, they share the thoughts and feelings of other customers who have previously used the hotel services. Thus, the more extreme the review sentiment for products and services is, the more helpful it is than neutral emotion. According to the above discussion, this study proposes the following hypotheses.

H2b. *Review sentiment will significantly affect the review's helpfulness.*

3.3. Moderating Variable: Luxury Hotel vs. Budget Hotel

Hotels can be divided into luxury hotels and budget hotels according to price and room size. When hotels are classified according to their star rating, four or more hotels are called luxury hotels [46,47]. In general, luxury hotel customers expect better hotel services (e.g., room size, information desk open time, foreign language-enabled staff, amenities, etc.) because they pay higher prices than in budget hotels.

Expectation mismatch theory is a theory developed by [22], in which customers are satisfied if the actual performance of a product or service is higher than expected and, conversely, dissatisfied if the actual performance is lower than expected. Therefore, hotel grades (luxury hotel vs. budget hotel) give customers different expectations for hotel services [48]. From this point of view, customers pay more attention to dissatisfaction reviews with the services of luxury hotels than to dissatisfaction reviews with the services of budget hotels and determine whether it helps to select hotels. Therefore, in this study, the normative and informational effects on review helpfulness will be moderated by the hotel grade (luxury hotel vs. budget hotel), and the hypothesis was established as follows.

H3a. *The effect of cleanliness on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H3b. *The effect of location on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H3c. *The effect of rooms on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H3d. *The effect of service on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H3e. *The effect of sleep on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H3f. *The effect of value on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H4a. *The effect of review length on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

H4b. *The effect of review sentiment on the review's helpfulness is more significant for ethnic hotels than budget hotels.*

4. Method

4.1. Data Collection

As of September 2021, this study collected review data from hotels in New York, Seoul, and Shanghai from TripAdvisor.com. TripAdvisor.com is the most popular travel information platform with more than 88 million reviews, and more than 900 million tourists use the platform to easily view others opinion of businesses, which can help with planning and booking of holidays.

The number of reviews collected is shown in Table 1. A total of 263,280 reviews were collected, including 227,661 reviews of 348 hotels in New York, 11,819 reviews of 228 hotels in Seoul, and 23,800 reviews of 431 hotels in Shanghai. In New York, there are approximately 654 reviews per hotel. In Seoul, there are approximately 52 reviews per hotel, and in Shanghai, there are approximately 55 reviews per hotel. It can be judged that New York in the United States has a higher average number of reviews than hotels in Seoul and Shanghai because it is a commercial, financial, economic, and cultural center.

Table 1. Number of reviews by city.

City	Number of Hotels	Number of Reviews
New York	348	227,661
Seoul	228	11,819
Shanghai	431	23,800
Sum	1007	263,280

Hotels with more than 4 stars are called luxury hotels [46,47], which are generally more expensive than budget hotels [47,49]. Therefore, customers expect a high level of service and facility quality for luxury hotels. Therefore, in this study, hotel ratings (hotel grades) were collected using hotel information from TripAdvisor.com. Table 2 shows the hotel ratings of the collected hotels. Approximately 63.9% in New York, 74.2% in Seoul, and 95.0% in Shanghai are luxury hotels.

Table 2. Number of reviews by hotel grades.

Hotel Grade	New York		Seoul		Shanghai	
	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
1	124	0.1%	3	0.0%	2	0.0%
1.5	3	0.0%	9	0.1%	0	0.0%
2	4266	1.9%	176	1.5%	119	0.5%
2.5	616	0.3%	42	0.4%	40	0.2%
3	69,573	30.6%	2719	23.0%	905	3.8%
3.5	7670	3.4%	99	0.8%	121	0.5%
4	121,192	53.2%	3226	27.3%	4507	18.9%
4.5	6680	2.9%	6	0.1%	1372	5.8%
5	17,537	7.7%	5539	46.9%	16,734	70.3%
Sum	227,661	100.0%	11,819	100.0%	23,800	100.0%

4.2. Operational Definition of a Variable

The variables for this study were extracted from the review of TripAdvisor.com, as described in Figure 2. The review shows the hotel’s cleanliness, location, rooms, service, sleep quality ratings, review text, and helpfulness votes.

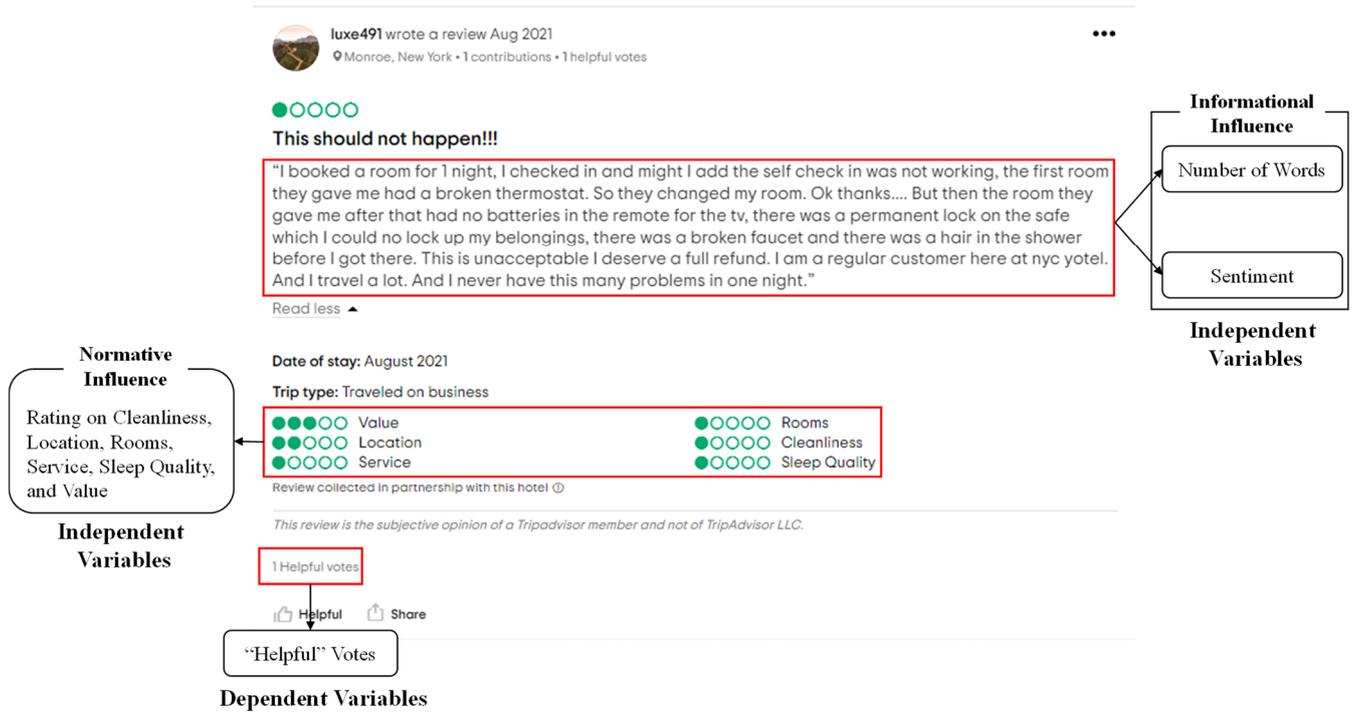


Figure 2. Illustration of the data variables.

The description of each variable is summarized in Table 3. Each variable was operationally defined based on reliable existing literature. As independent variables, cleanliness, location, rooms, service, sleep quality, value, review length, and review emotion shown in the review were used. The rating of cleanliness, location, rooms, service, sleep quality, and value was extracted directly from the review data. The review length was extracted from the review using the LIWC program. Review sentiments were extracted using the Python program’s VADER (valence aware dictionary and sentimental reasoner) package. Helpfulness was used as a dependent variable and was extracted directly from the review data. Finally, the moderator variable, the hotel grade, was coded as 1 for luxury hotels (when the hotel grade is 4 stars or more) and 0 for other cases.

Table 3. Operationalization of variables.

Type	Variables	Operationalization	Notes	Reference
Independent Variable	Cleanliness	The review’s cleanliness rating from TripAdvisor	Range: [0, 5]	[50,51]
	Location	The review’s location rating from TripAdvisor	Range: [0, 5]	[50,51]
	Rooms	The review’s rooms rating from TripAdvisor	Range: [0, 5]	[50,51]
	Service	The review’s service rating from TripAdvisor	Range: [0, 5]	[50,51]
	Sleep Quality	The review’s sleep quality rating from TripAdvisor	Range: [0, 5]	[50,51]
	Value	The review’s value rating from TripAdvisor	Range: [0, 5]	[50,51]
Informational Influence	Review Length	Number of words in the review	Continuous	[9,13,52]
	Sentiment	The absolute value of the sentiment analysis score extracted from the review using Python’s VADER package	range [0, 1]	[15,43]
Moderating Variable	Hotel Grade	Hotel grade of 4 or higher (luxury) is 1; otherwise (budget) is 0	Range: [0, 1]	[50,53]
Dependent Variable	Helpfulness	Number of helpfulness votes received by the review	Continuous	[9,16]

4.3. Descriptive Analysis

The descriptive statistics and correlation analysis for the variables in this study are shown in Table 4. The ratings for 6 attributes of hotels in each city is 4 or higher, indicating that customers are overall satisfied with the hotel. The review sentiment is 0.8 or higher, indicating that customers had a positive hotel experience. However, the number of helpful votes is very low, approximately 0.2 per review. This means that most reviews are not helpful. Furthermore, there are relatively high correlations between some variables, so multicollinearity is likely to occur.

Table 4. Descriptive analysis for the variables.

City	Variables	Min	Max	Mean	Standard Deviation	1	2	3	4	5	6	7	8	9
All Cities	1. Cleanliness	1	5	4.43	0.89	1								
	2. Location	1	5	4.64	0.68	0.409 ***	1							
	3. Rooms	1	5	4.15	1.03	0.730 ***	0.418 ***	1						
	4. Service	1	5	4.32	1.03	0.662 ***	0.420 ***	0.654 ***	1					
	5. Sleep Quality	1	5	4.29	1.00	0.654 ***	0.409 ***	0.715 ***	0.607 ***	1				
	6. Value	1	5	4.07	1.06	0.648 ***	0.427 ***	0.707 ***	0.690 ***	0.638 ***	1			
	7. Review length	6	4332	113.90	109.71	-0.087 ***	-0.081 ***	-0.098 ***	-0.155 ***	-0.106 ***	-0.109 ***	1		
	8. Sentiment	0	1	0.86	0.20	0.338 ***	0.218 ***	0.361 ***	0.368 ***	0.340 ***	0.371 ***	0.163 ***	1	
	9. Helpfulness	0	137	0.16	0.56	-0.016 ***	0.016 ***	-0.011 ***	-0.008 ***	-0.009 ***	-0.016 ***	-0.089 ***	-0.049 ***	1
New York	1. Cleanliness	1	5	4.43	0.90	1								
	2. Location	1	5	4.68	0.64	0.429 ***	1							
	3. Rooms	1	5	4.13	1.04	0.731 ***	0.447 ***	1						
	4. Service	1	5	4.33	1.03	0.668 ***	0.431 ***	0.663 ***	1					
	5. Sleep Quality	1	5	4.28	1.02	0.652 ***	0.434 ***	0.716 ***	0.612 ***	1				
	6. Value	1	5	4.07	1.08	0.657 ***	0.446 ***	0.717 ***	0.697 ***	0.643 ***	1			
	7. Review length	6	4332	110.65	106.00	-0.091 ***	-0.066 ***	-0.105 ***	-0.156 ***	-0.111 ***	-0.114 ***	1		
	8. Sentiment	0	1	0.87	0.20	0.345 ***	0.227 ***	0.371 ***	0.376 ***	0.348 ***	0.382 ***	0.159 ***	1	
	9. Helpfulness	0	137	0.17	0.59	-0.016 ***	0.005 *	-0.009 ***	-0.012 ***	-0.008 ***	-0.018 ***	-0.093 ***	-0.052 ***	1
Seoul	1. Cleanliness	1	5	4.43	0.83	1								
	2. Location	1	5	4.36	0.85	0.367 ***	1							
	3. Rooms	1	5	4.16	0.94	0.686 ***	0.379 ***	1						
	4. Service	1	5	4.31	0.94	0.608 ***	0.390 ***	0.612 ***	1					
	5. Sleep Quality	1	5	4.27	0.92	0.649 ***	0.358 ***	0.692 ***	0.578 ***	1				
	6. Value	1	5	4.04	0.96	0.552 ***	0.440 ***	0.623 ***	0.617 ***	0.578 ***	1			
	7. Review length	8	2755	140.84	138.88	-0.065 ***	-0.070 ***	-0.097 ***	-0.121 ***	-0.104 ***	-0.082 ***	1		
	8. Sentiment	0	1	0.86	0.20	0.263 ***	0.182 ***	0.273 ***	0.289 ***	0.256 ***	0.269 ***	0.208 ***	1	
	9. Helpfulness	0	5	0.12	0.42	0.012	0.030 ***	0.039 ***	0.020 *	0.026 **	0.030 ***	-0.063 ***	-0.020 *	1
Shanghai	1. Cleanliness	1	5	4.47	0.83	1								
	2. Location	1	5	4.34	0.88	0.381 ***	1							
	3. Rooms	1	5	4.35	0.88	0.750 ***	0.411 ***	1						
	4. Service	1	5	4.22	1.04	0.637 ***	0.401 ***	0.634 ***	1					
	5. Sleep Quality	1	5	4.40	0.88	0.681 ***	0.389 ***	0.710 ***	0.596 ***	1				
	6. Value	1	5	4.14	0.94	0.597 ***	0.401 ***	0.635 ***	0.659 ***	0.602 ***	1			
	7. Review length	8	2143	131.66	123.75	-0.083 ***	-0.090 ***	-0.079 ***	-0.163 ***	-0.091 ***	-0.100 ***	1		
	8. Sentiment	0	1	0.86	0.20	0.298 ***	0.206 ***	0.313 ***	0.331 ***	0.300 ***	0.309 ***	0.179 ***	1	
	9. Helpfulness	0	4	0.06	0.29	-0.013 *	0.014 *	0.000	-0.005	0.001	-0.001	-0.022 ***	-0.018 **	1

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4.4. Analyze Data and Test Hypotheses

In this study, the distribution of dependent variables to select a regression analysis model is as follows. First, helpfulness, a dependent variable, is a discrete (countable) variable. Second, the variance of helpfulness is greater than the mean (see Table 4). In general, negative binomial regression is used when the variance of the dependent variable is greater than the mean [54]. Finally, the helpfulness distribution has many extremes of 0, as shown in Figure 3.

Therefore, ZINB (zero-inflated negative binomial regression) was used in this study. In addition, in this study, a 4-step hierarchical regression analysis was conducted to analyze the hotel-grade interaction effect on normative influencing factors (cleanness, location, rooms, service, sleep quality, value) and informational influencing factors (review length, review sentiment). In Step 1, the control variable (city) was considered. In Step 2, the independent variable (cleanness, location, rooms, service, sleep quality, value, review length, and review sentiment) and the moderating variable (hotel grade) were considered. In addition, the interaction term for normative effects was included in Step 3, and the interaction term for informational effects was added in Step 4.

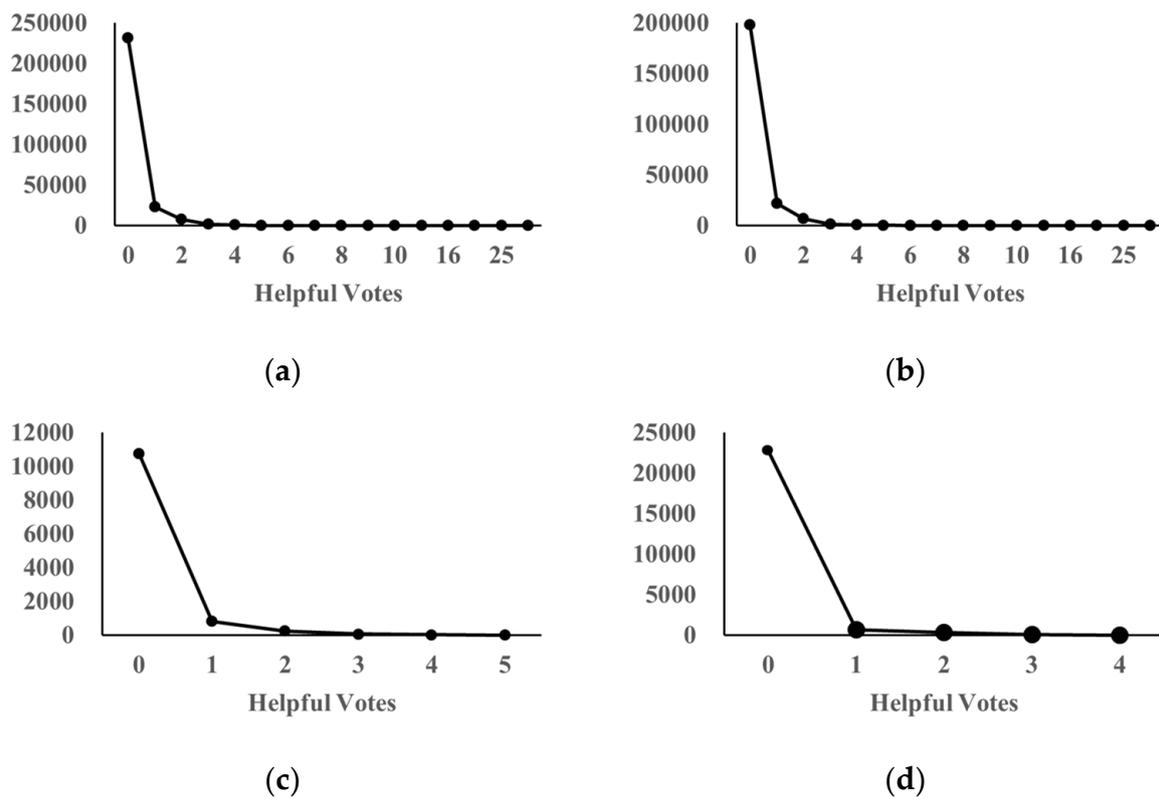


Figure 3. Distribution of Helpfulness by City. (a) All Cities; (b) New York; (c) Seoul; (d) Shanghai.

When analyzing the hierarchical moderating effect, the interaction term between the independent variable and the moderating variable can cause multicollinearity between the independent variable and the moderating variable, so VIF (variance inflation factor) was confirmed. As a result, since the VIF is greater than 10, it is performed mean-centering on the independent and moderating variables to minimize the potential problem of multicollinearity. As a result, the VIFs ranged from 1 to 3.2, significantly lower than 10. Therefore, the multicollinearity problem no longer exists.

Model 1 reflects the relationship between helpfulness, the normative influencing factors (cleanness, location, rooms, service, sleep quality, value), and the informational influencing factors (review length, review sentiment), as mentioned in H1a, H1b, H1b, H1a, H1b, H2b, H3c, H3d, H3e, H3f, H4a, and H3b.

$$\text{Model 1: Helpfulness} = \beta_{11} \times \text{Cleanliness} + \beta_{12} \times \text{Location} + \beta_{13} \times \text{Rooms} + \beta_{14} \times \text{Service} + \beta_{15} \times \text{Sleep Quality} + \beta_{16} \times \text{Value} + \beta_{17} \times \text{Number of Words} + \beta_{18} \times \text{Sentiment} + \beta_{19} \times \text{Cleanliness} \times \text{Hotel Grade} + \beta_{20} \times \text{Location} \times \text{Hotel Grade} + \beta_{21} \times \text{Rooms} \times \text{Hotel Grade} + \beta_{22} \times \text{Service} \times \text{Hotel Grade} + \beta_{23} \times \text{Sleep Quality} \times \text{Hotel Grade} + \beta_{24} \times \text{value} \times \text{Hotel Grade} + \beta_{25} \times \text{Review Length} \times \text{Hotel Grade} + \beta_{26} \times \text{Sentiment} \times \text{Hotel Grade} + D_{\text{NewYork}} + D_{\text{Shanghai}} + \varepsilon_1$$

Table 5 shows the negative binomial regression analysis results for all hotels regardless of city. The effects of normative influencing factors (cleanness, location, rooms, service, sleep quality, value) and informational influencing factors (review length, review sentiment) on review helpfulness were analyzed.

Table 5. Results of zero-inflated negative binomial analysis (all cities).

Variables	Step 1			Step 2			Step 3			Step 4		
	Coef.	Std. er.	z-Value									
(Constant)	0.231	0.028	8.40	0.394	0.050	7.89	0.404	0.050	8.08	0.390	0.050	7.77
Step 1: City (ref.: Seoul)												
Shanghai	0.031	0.040	0.77	0.019	0.040	0.48	0.016	0.040	0.39	0.008	0.040	0.21
New York	−0.489	0.029	−17.14 ***	−0.507	0.029	−17.59 ***	−0.517	0.029	−17.86 ***	−0.519	0.029	−17.94 ***
Step 2: Main Effect (Independent and Moderator Variable)												
Cleanliness				−0.021	0.010	−2.20 *	−0.020	0.010	−2.14 *	−0.020	0.010	−2.09 *
Location				0.038	0.010	3.77 ***	0.037	0.010	3.62 ***	0.037	0.010	3.66 ***
Rooms				0.016	0.010	1.60	0.017	0.010	1.70	0.015	0.010	1.54
Service				−0.026	0.008	−3.20 **	−0.026	0.008	−3.15 ***	−0.026	0.008	−3.20 *
Sleep Quality				−0.001	0.008	−0.08	−0.001	0.008	−0.18 *	−0.002	0.008	−0.18 **
Value				−0.040	0.009	−4.56 ***	−0.039	0.009	−4.46 ***	−0.038	0.009	−4.32
Review Length				−0.000	0.000	−3.98 ***	0.000	−0.000	−4.050 ***	−0.000	0.000	−4.74 **
Sentiment				−0.039	0.029	−1.36	−0.038	0.029	−1.33	−0.038	0.029	−1.31
Hotel Grade				0.049	0.012	4.04 ***	0.044	0.012	3.58 ***	0.073	0.014	5.17 ***
Step 3: Hotel-grade interaction effects on normative influencing factors												
Cleanliness × Hotel Grade							−0.061	0.020	−2.97 ***	−0.064	0.020	−3.13 **
Location × Hotel Grade							0.060	0.021	2.78 **	0.060	0.021	2.81 **
Rooms × Hotel Grade							−0.016	0.021	−0.75	−0.014	0.021	−0.66
Service × Hotel Grade							−0.007	0.018	−0.40	−0.001	0.018	−0.04
Sleep Quality × Hotel Grade							0.026	0.018	1.49	0.027	0.018	1.53
Value × Hotel Grade							−0.008	0.019	−0.44	−0.011	0.019	−0.57
Step 4: Hotel-grade interaction effects on informational influencing factors												
Review Length × Hotel Grade										0.001	0.000	4.06 ***
Sentiment × Hotel Grade										0.031	0.062	0.50
Log Likelihood		−100,949.8			−100,863.2			−100,850.1			−100,840.9	
LR Chi-square		507.37 ***			680.55 ***			706.91			725.30 ***	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

The analysis results are as follows. In the case of New York hotels, it was found that they had a negative (−) effect on the review's helpfulness compared to Seoul hotels. In addition, as assumed in this study, the normative factors of cleanliness, location, service, and value, the informational influencing factor of review length, and the moderator variable, hotel grade, significantly affect the review's helpfulness in Steps 2 to 4.

In Step 2, from the perspective of the normative influence, the high rating of the location ($\beta = 0.038$, $p < 0.001$) has a significantly positive (+) effect on the review's helpfulness. However, the high ratings of cleanliness ($\beta = -0.021$, $p < 0.05$), service ($\beta = -0.026$, $p < 0.01$), and value ($\beta = -0.040$, $p < 0.001$) have significant negative (−) effects on the review's helpfulness. From the perspective of the informational influence, the longer the review length ($\beta = -0.000$, $p < 0.001$) is, the more significantly the negative (−) effect is. Hotels with a high hotel grade ($\beta = 0.049$, $p < 0.001$), a moderator variable, have a significantly positive (+) effect on the review's helpfulness. Therefore, this study supports the hypotheses H1a, H1b, H1d, H1f, and H2a. Location × hotel grade interaction ($\beta = 0.060$, $p < 0.01$) has a significantly positive (+) effect on the review's helpfulness in Step 3 and Step 4. On the other hand, the cleanliness × hotel grade interaction ($\beta = -0.061$, $p < 0.001$) has a significantly negative (−) effect on the review's helpfulness in Steps 3 and 4. Moreover, the review length × hotel grade ($\beta = 0.001$, $p < 0.001$) has a significant positive (+) effect on the review's helpfulness in Step 4. Accordingly, H3a, H3b, and H4a are supported, but H3c, H3d, H3e, H3f, and H4b are not supported.

In addition, in this study, a 3-step hierarchical regression analysis was conducted to analyze the hotel-grade interaction effect on normative influencing factors (cleanness, location, rooms, service, sleep quality, and value) and informational influencing factors (review length, and review sentiment). In Step 1, independent variables (cleanness, location, rooms, service, sleep quality, value, review length, and review sentiment) and moderator variables (hotel grade) were considered. In Step 2, interaction terms for normative effects were included. Finally, in Step 3, an interaction term for informational effects was included.

Model 2 reflects the hypothesis relationship between normative influencing factors, informational influencing factors, and review's helpfulness, as mentioned in Model 1.

$$\text{Model 2: Helpfulness} = \beta_{11} \times \text{Cleanliness} + \beta_{12} \times \text{Location} + \beta_{13} \times \text{Rooms} + \beta_{14} \times \text{Service} + \beta_{15} \times \text{Sleep Quality} + \beta_{16} \times \text{Value} + \beta_{17} \times \text{Number of Words} + \beta_{18} \times \text{Sentiment} + \beta_{19} \times \text{Cleanliness} \times \text{Hotel Grade} + \beta_{20} \times \text{Location} \times \text{Hotel Grade} + \beta_{21} \times \text{Rooms} \times \text{Hotel Grade} + \beta_{22} \times \text{Service} \times \text{Hotel Grade} + \beta_{23} \times \text{Sleep Quality} \times \text{Hotel Grade} + \beta_{24} \times \text{value} \times \text{Hotel Grade} + \beta_{25} \times \text{Review Length} \times \text{Hotel Grade} + \beta_{26} \times \text{Sentiment} \times \text{Hotel Grade} + \varepsilon_1$$

First, the results of the negative binomial regression analysis for New York hotels are shown in Table 6. Normative influencing factors including cleanliness, location, rooms, service, and value, informational influencing factors including review length and review sentiment among, and the moderator variable, hotel grade, significantly affect the review's helpfulness in Steps 1 to 3.

Table 6. Results of zero-inflated negative binomial analysis (New York).

Variables	Step 1			Step 2			Step 3		
	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value
(Constant)	−0.689	0.050	−13.90	−0.695	0.050	−14.03	−0.713	0.050	−14.36
Step 1: Main Effect (Independent and Moderator Variable)									
Cleanliness	−0.022	0.010	−2.09 *	−0.020	0.011	−1.92	−0.020	0.011	−1.89
Location	0.069	0.011	6.07 ***	0.069	0.011	6.10 ***	0.069	0.011	6.07 ***
Rooms	0.024	0.010	2.34 *	0.026	0.010	2.44 *	0.024	0.010	2.32 *
Service	−0.040	0.009	−4.39 ***	−0.040	0.009	−4.45 ***	−0.041	0.009	−4.45 ***
Sleep Quality	−0.015	0.009	−1.66	−0.016	0.009	−1.76	−0.016	0.009	−1.81
Value	−0.050	0.009	−5.33 ***	−0.050	0.009	−5.36 ***	−0.048	0.009	−5.16 ***
Review Length	−0.004	0.000	−42.68 ***	−0.004	0.000	−42.60 ***	−0.004	0.000	−42.74 ***
Sentiment	−0.252	0.032	−7.96 ***	−0.250	0.032	−7.90 ***	−0.249	0.032	−7.89 ***
Hotel Grade	0.140	0.013	10.82 ***	0.136	0.013	10.50 ***	0.170	0.015	11.62 ***
Step 2: Hotel-grade interaction effects on normative influencing factors									
Cleanliness × Hotel Grade				−0.057	0.022	−2.59 *	−0.062	0.022	−2.77 **
Location × Hotel Grade				0.093	0.023	3.99 ***	0.092	0.023	3.95 ***
Rooms × Hotel Grade				−0.057	0.022	−2.59 **	−0.055	0.022	−2.50 *
Service × Hotel Grade				−0.008	0.019	−0.39	0.000	0.019	−0.02
Sleep Quality × Hotel Grade				0.015	0.019	0.77	0.016	0.019	0.86
Value × Hotel Grade				0.049	0.020	2.47 *	0.046	0.020	2.31 *
Step 3: Hotel-grade interaction effects on informational influencing factors									
Review Length × Hotel Grade							0.001	0.000	4.88 ***
Sentiment × Hotel Grade							0.032	0.066	0.48
Log Likelihood		−105,246.6			−105,228.9			−105,215.6	
LR Chi-square		2853.85 ***			2889.39 ***			2915.93 ***	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

In Step 1, from the perspective of the normative influence, the high rating of the location ($\beta = 0.069$, $p < 0.001$) and rooms ($\beta = 0.024$, $p < 0.05$) have a significantly positive (+) effect on the review's helpfulness. However, the high ratings of cleanliness ($\beta = -0.022$, $p < 0.05$), service ($\beta = -0.040$, $p < 0.001$), and value ($\beta = -0.050$, $p < 0.001$) have significant negative (−) effects on the review's helpfulness. From the perspective of the informational influence, the longer the review length ($\beta = -0.004$, $p < 0.001$) is, the more significant the negative (−) effect is. Moreover, the more extreme reviews have a significant negative (−) effect on the review's helpfulness. Hotels with a high hotel grade ($\beta = 0.140$, $p < 0.001$), a moderator variable, have a significantly positive (+) effect on the review's helpfulness. Therefore, this study supports hypotheses H1a, H1b, H1c, H1d, H1f, H2a, and H2b. Location × hotel grade interaction ($\beta = 0.093$, $p < 0.001$) and value × hotel grade interaction ($\beta = 0.049$, $p < 0.05$) have a significantly positive (+) effect on the review's helpfulness in Step 2 and Step 3. On the other hand, the cleanliness × hotel grade interaction ($\beta = -0.057$, $p < 0.05$) and rooms × hotel grade interaction have a significantly negative (−) effect on the review's helpfulness in steps 2 and 3. Moreover, the review length × hotel grade ($\beta = 0.001$, $p < 0.001$) has a significantly positive (+) effect on the review's helpfulness in Step 3. Accordingly, H3a, H3b, H3c, H3f, and H4a are supported, but H3d, H3e, and H4b are not supported.

Second, the results of the negative binomial regression analysis for Seoul hotels are shown in Table 7. Location in normative influencing factors and review length in informational influencing factors significantly affect the review's helpfulness in Steps 1 to 3.

Table 7. Results of zero-inflated negative binomial analysis (Seoul).

Variables	Step 1			Step 2			Step 3		
	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value
(Constant)	−0.805	0.223	−3.60	−0.881	0.236	−3.73	−0.878	0.236	−3.72
Step 1: Main Effect (Independent and Moderator Variable)									
Cleanliness	−0.079	0.059	−1.35	−0.063	0.059	−1.06	−0.064	0.059	−1.09
Location	0.131	0.045	2.90 **	0.133	0.047	2.84 **	0.132	0.047	2.82 **
Rooms	0.085	0.059	1.43	0.073	0.059	1.23	0.075	0.059	1.26
Service	−0.088	0.048	−1.83	−0.066	0.049	−1.35	−0.064	0.049	−1.30
Sleep Quality	−0.034	0.053	−0.65	−0.042	0.052	−0.79	−0.041	0.052	−0.78
Value	0.065	0.050	1.31	0.044	0.051	0.88	0.043	0.051	0.84
Review Length	−0.001	0.000	−2.26 *	−0.001	0.000	−2.29 *	−0.001	0.000	−2.29 *
Sentiment	−0.161	0.163	−0.99	−0.141	0.162	−0.87	−0.139	0.163	−0.86
Hotel Grade	0.117	0.085	1.37	0.144	0.094	1.54	0.132	0.094	1.40
Step 2: Hotel-grade interaction effects on normative influencing factors									
Cleanliness × Hotel Grade				0.190	0.128	1.49	0.199	0.128	1.55
Location × Hotel Grade				−0.141	0.123	−1.15	−0.143	0.123	−1.17
Rooms × Hotel Grade				0.038	0.139	0.27	0.033	0.140	0.23
Service × Hotel Grade				−0.267	0.125	−2.13 *	−0.283	0.126	−2.24 *
Sleep Quality × Hotel Grade				−0.065	0.123	−0.53	−0.068	0.123	−0.56
Value × Hotel Grade				0.353	0.127	2.78 **	0.356	0.127	2.80 **
Step 3: Hotel-grade interaction effects on informational influencing factors									
Review Length × Hotel Grade							−0.001	0.001	−1.07
Sentiment × Hotel Grade							−0.020	0.385	−0.05
Log Likelihood		−4156.68			−4148.868			−4148.271	
LR Chi-square		30.88 ***			42.95 ***			44.14 ***	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

In Step 1, from the perspective of normative influence, the high rating of the location ($\beta = 0.131$, $p < 0.01$) has a significantly positive (+) effect on the review's helpfulness. From the perspective of informational influence, the longer the review length ($\beta = -0.001$, $p < 0.05$) is, the more significant the negative (−) effect is. Therefore, this study supports only hypotheses H1a and H2a. On the other hand, the value × hotel grade interaction ($\beta = 0.093$, $p < 0.01$) has a significantly positive (+) effect on the review's helpfulness in Step 2 and Step 3, and the service × hotel grade interaction ($\beta = -0.267$, $p < 0.05$) has a significantly negative (−) effect on the review's helpfulness in Steps 2 and 3. Accordingly, H3d and H3f are supported, but the others are not supported.

Third, the results of the negative binomial regression analysis for Shanghai hotels are shown in Table 8. Only cleanliness in normative influencing factors significantly affects the review's helpfulness in Steps 1 to 3.

In Step 1, from the perspective of the normative influence, the high rating of the cleanliness ($\beta = 0.131$, $p < 0.01$) has a significantly negative (−) effect on the review's helpfulness. Moreover, any interaction of normative and informational influence with hotel grade does not affect the review's helpfulness. Therefore, this study supports only hypothesis H1a.

Table 8. Results of zero-inflated negative binomial analysis (Shanghai).

Variables	Step 1			Step 2			Step 3		
	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value	Coef.	Std. er.	z-Value
(Constant)	−0.805	0.223	−3.60	0.106	0.271	0.39	0.099	0.273	0.36
Step 1: Main Effect (Independent and Moderator Variable)									
Cleanliness	−0.177	0.060	−2.98 **	−0.189	0.060	−3.18 **	−0.189	0.060	−3.18 **
Location	0.084	0.045	1.85	0.084	0.046	1.85	0.084	0.046	1.85
Rooms	0.045	0.061	0.74	0.038	0.061	0.62	0.038	0.061	0.62
Service	−0.063	0.047	−1.33	−0.061	0.047	−1.31	−0.061	0.047	−1.30
Sleep Quality	0.023	0.057	0.41	0.031	0.057	0.54	0.031	0.057	0.54
Value	0.089	0.053	1.69	0.096	0.053	1.83	0.097	0.053	1.83
Review Length	0.000	0.000	−1.38	0.000	0.000	−1.39	0.000	0.000	−1.38
Sentiment	−0.238	0.173	−1.37	−0.237	0.173	−1.37	−0.237	0.173	−1.37
Hotel Grade	−0.117	0.158	−0.74	−0.272	0.166	−1.64	−0.266	0.170	−1.57
Step 2: Hotel-grade interaction effects on normative influencing factors									
Cleanliness × Hotel Grade				−0.288	0.243	−1.18	−0.286	0.246	−1.16
Location × Hotel Grade				−0.052	0.187	−0.28	−0.049	0.188	−0.26
Rooms × Hotel Grade				−0.123	0.306	−0.40	−0.126	0.325	−0.39
Service × Hotel Grade				−0.225	0.224	−1.01	−0.223	0.227	−0.98
Sleep Quality × Hotel Grade				0.085	0.184	0.46	0.086	0.184	0.47
Value × Hotel Grade				0.253	0.245	1.03	0.248	0.248	1.00
Step 3: Hotel-grade interaction effects on informational influencing factors									
Review Length × Hotel Grade							0.000	0.001	0.17
Sentiment × Hotel Grade							0.025	0.829	0.03
Log Likelihood		−4751.558			−4747.432			−4747.414	
LR Chi-square		23.74 **			31.99 ***			32.03 *	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4.5. Robustness Checking

In this study, we verified the robustness of the proposed method with two regression analyses. The first is a Tobit regression analysis. The Tobit regression analysis, used in many studies, was used to check the robustness of the method [9,55]. The second is 2-stage least squares (2SLS). 2SLS is one of the methods used to solve the endogeneity problem [56,57]. There may be unabsorbed variables that influence both review length and review helpfulness. For example, some reviewers may like to write in detail about the hotel services they used, while others may not. Therefore, short reviews may be less likely to receive helpful votes. We used average review length as an instrument which provides an exogenous source of variation for the endogenous variable (review length).

The robustness check results are shown in Table 9. There was almost no significant difference between the results of the ZINB, the Tobit regression analysis, and the 2SLS, and the main effect remained similar. Therefore, it can be concluded that the analysis results of the proposed research model in this study are generally robust.

Table 9. Robustness check results for alternative model specification.

Variables	All Cities			New York			Seoul			Shanghai		
	ZINB	Tobit	2SLS	ZINB	Tobit	2SLS	ZINB	Tobit	2SLS	ZINB	Tobit	2SLS
Cleanliness	−0.020 *	−0.052 ***	−0.009 ***	−0.020	−0.042 **	−0.009 ***	−0.064	−0.219	−0.013	−0.189 **	−0.258 **	−0.013 ***
Location	0.037 ***	0.213 ***	0.024 ***	0.069 ***	0.122 **	0.016 ***	0.132 **	0.138	0.012 *	0.084	0.155 **	0.007 **
Rooms	0.015	0.010	0.002	0.024 *	0.039 **	0.005 **	0.075	0.180	0.015 *	0.038	0.128	0.005
Service	−0.026 *	−0.029 *	−0.003	−0.041 ***	−0.051 ***	−0.006 ***	−0.064	−0.019	−0.007	−0.061	−0.055	−0.003
Sleep Quality	−0.002 **	−0.006	−0.001	−0.016	0.006	0.000	−0.041	−0.009	−0.002	0.031	0.049	0.004
Value	−0.038	−0.078 ***	−0.009 ***	−0.048 ***	−0.068 ***	−0.008 ***	0.043	0.097	0.010	0.097	−0.012	0.001
Review Length	−0.000 **	−0.007 ***	0.000 ***	−0.004 ***	−0.007 ***	−0.001 ***	−0.001 *	−0.002	0.000 ***	0.000	−0.001 **	0.000 ***
Sentiment	−0.038	−0.566 ***	−0.077 ***	−0.249 ***	−0.598 ***	−0.087 ***	−0.139	−0.497	−0.038	−0.237	−0.425	−0.021 *
Hotel Grade	0.073 ***	0.085 ***	0.002	0.170 ***	0.199 ***	0.015 ***	0.132	0.433	0.043 ***	−0.266	−0.561 *	−0.023
Cleanliness × Hotel Grade	−0.064 **	−0.092 **	−0.014 ***	−0.062 **	−0.089 **	−0.015 ***	0.199	0.211	0.019	−0.286	−0.445	−0.022
Location × Hotel Grade	0.060 **	0.131 ***	0.015 ***	0.092 ***	0.045	0.008	−0.143	−0.271	−0.018	−0.049	−0.269	−0.006
Rooms × Hotel Grade	−0.014	−0.216 ***	−0.025 **	−0.055 *	−0.180 ***	−0.021 ***	0.033	0.093	0.009	−0.126	−0.074	−0.006
Service × Hotel Grade	−0.001	−0.095 ***	−0.010 **	0.000	−0.098 ***	−0.012 **	−0.283 *	−0.446	−0.036 **	−0.223	−0.530	−0.019
Sleep Quality × Hotel Grade	0.027	−0.009	0.000	0.016	0.017	0.002	−0.068	−0.244	−0.015	0.086	−0.183	0.003
Value × Hotel Grade	−0.011	0.220 ***	0.023 ***	0.046 *	0.207 ***	0.022 **	0.356 **	0.656	0.051 ***	0.248	0.506	0.018
Review Length × Hotel Grade	0.001 ***	0.002 ***	0.000 ***	0.001 ***	0.002 ***	0.000	−0.001	−0.002	0.000 *	0.000	0.001	0.000
Sentiment × Hotel Grade	0.031	0.134	0.016	0.032	0.117	0.012	−0.020	−0.139	−0.022	0.025	−0.292	−0.020

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4.6. Discussion

This study investigates the effects of normative and informational influencing factors on review helpfulness based on the social influence theory. First, our study found that normative influencing factors (i.e., cleanliness, location, service, sleep quality) and informative influencing factors (i.e., review length) affect the helpfulness of reviews, regardless of the city's location. This is consistent with the social influence theory, where customers are influenced by other customers' human behavior [58]. Furthermore, the effect of normative and informational influencing factors on review helpfulness is moderated by hotel ratings (i.e., luxury hotel vs. budget hotel). This is consistent with the expectation disconfirmation theory, where the higher the price, the better the service customers expect. Therefore, we employed the social influence theory and the expectation disconfirmation theory as a theoretical basis to explain the helpfulness of online reviews for hotels. In particular, among the normative influencing factors, cleanliness, service, and sleep quality were found to have a negative effect on review helpfulness. In fact, according to the signal theory, customers cannot accurately know the service quality of a hotel due to information asymmetry. Therefore, review rating is a good signal to resolve the uncertainty regarding the service quality of the hotel. As negative review ratings mean they are dissatisfied with the quality of hotel service [59], this is an important signal, especially for customers who think that a particular hotel's service is better than other accommodations. Therefore, the results of this study can be said to be consistent with the results of previous studies where cleanliness, service, and quality of sleep have a negative effect on reviews [60].

Second, our findings revealed that there were differences between normative and informational influencing factors affecting review helpfulness according to the geographic location of hotels. For example, in New York, the determinants of the review's helpfulness are cleanliness, location, rooms, service, value, review length, and review sentiment. The location and review length for Seoul hotels and cleanliness for Shanghai hotels are the determinants of the review's helpfulness.

Third, in this study, the moderating effect of hotel grade was confirmed in the relationship between normative influencing factors, informational influencing factors, and the review's helpfulness. In general, the higher the price of a hotel, the higher the expectation in both luxury hotels and budget hotels. In New York, when analyzing the more expensive hotels compared with budget hotels, the lower the cleanliness and room ratings, the higher the review's helpfulness. However, higher location and value ratings and longer review lengths have a negative effect on review helpfulness.

Finally, many previous studies have argued that a longer review length has a positive effect on review helpfulness because the information in the review is richer [10,13,14,31,32,35–39]. However, in this study, review length was found to have a negative effect on review helpfulness. In general, if a problem is not related to the consumer's interest, the motivation for cognitive effort does not occur, but if it is related to profit, the motivation for cognitive effort is apparent [52]. Therefore, when a customer has a negative experience, they actively express their opinion in the review to help others make decisions. Therefore, the longer the review length, the more negative opinions can be, and it can be inferred in which part the negative experience was made through the normative influence factors.

5. Conclusions

5.1. Theoretical Implications

The theoretical contributions of this study are as follows. First, previous studies have explained the review rating, which indicates the overall satisfaction of a hotel, as a normative influencing factor. However, our study examined the influence of multiple attributes, which indicate the detailed service of the hotel as a normative influencing factor, on review helpfulness. As a result of our study, cleanliness, location, service, and sleep quality were found to affect review helpfulness. Therefore, our study contributes to future online review research by extending the normative influencing factors in the online review literature.

Second, although the customer's evaluation of a hotel differs according to the geographical location of the hotel, previous studies investigated the determinants of review helpfulness for hotels in a single city. Accordingly, this study collected reviews of hotels in New York, Seoul, and Shanghai, and investigated factors affecting the helpfulness of reviews. As a result, it was revealed that the determinants of review helpfulness for hotels in each city were different. Therefore, our study contributes to the online review literature by investigating the determinants of review helpfulness according to the geographic location of the hotel.

Finally, unlike previous studies which analyzed the influence of review ratings on helpfulness according to hotel grade (luxury hotel vs. budget hotel), our study analyzed the influence of multiple attributes on helpfulness according to hotel grade. Therefore, our study contributes to the online review literature by examining the influence of hotel service attributes on review helpfulness according to hotel grade.

5.2. Practical Implications

The practical implications derived from the results of this study are as follows. First, the results of this study can be used as a guideline for predicting meaningful reviews. Hotel managers will help improve negative reviews and enhance the promotion of positive reviews, potentially helping to enhance customer satisfaction. Second, identifying the factors that affect the review's helpfulness will help strengthen customer satisfaction. In other words, it will help hotel managers take improvement measures on factors that negatively affect the review's helpfulness and strengthen factors that positively affect the review's helpfulness. Therefore, it will potentially help to improve the hotel's profitability. Third, approximately 14% of hotel ratings are affected by travel destinations [61]. In addition, the factors affecting the review's helpfulness vary depending on the hotel's geographical location. Therefore, in the case of franchise hotels, it is necessary to provide locally differentiated services. Fourth, the evaluation of the detailed elements of the review reflects the difference in the performance of customers' expectations and experiences at the hotel quality level. In other words, if the rating is high, the performance is higher than expected, while if the rating is low, the performance is lower than expected. Customers expect better hotel quality the more expensive the hotel is. Therefore, in luxury hotels, the lower the detailed rating of the hotel, the more helpful the review tends to be. Since this can negatively affect customers' purchase of luxury hotels, luxury hotel managers must improve the service quality dimension of hotels with low ratings.

5.3. Recommendation, Limitations, and Future Research

As a result of this study, it was found that the negative multi-attribute ratings for hotel service affect the helpfulness of reviews. In addition, it was found that there was a difference in the influence of the hotel's multi-attribute ratings on the helpfulness according to the hotel grade. On this base, future research should develop evaluation criteria of multiple attributes for hotel service and a methodology to evaluate it.

However, this study has the following limitations. First, the evaluation of hotels has different expectations for each customer due to cultural differences. In other words, customers with different cultural backgrounds may have different expectations and perceptions of hotel services. However, since this study only collected English reviews from hotels in New York, Seoul, and Shanghai to analyze the determinants that affect the review's helpfulness, it is necessary to collect and analyze reviews in various languages, such as Korean and Chinese. Second, since hotel review data were collected from a single travel platform (TripAdvisor.com), there is a limit to generalizing the research results. In future studies, collecting and analyzing data from various travel platforms such as Expedia.com and Booking.com is necessary. Third, in many studies, the determinants of the review's helpfulness were identified using the review's attribute data and the reviewer's attribute data. However, in this study, it was analyzed using only the attribute data of the review. In future studies, it is necessary to add and analyze the reviewer's attribute data. Finally, in

this study, only the data of the ranking top-1000 provided by TripAdvisor were collected, so the proportion of luxury hotels was high. In China, most of the data used in the analysis were from luxury hotels, so there was a limit to being unable to analyze the moderating effect of the hotel grade. Therefore, in future studies, it is necessary to collect and analyze data by expanding the scope of hotels.

Author Contributions: Conceptualization, Y.S. and I.C.; methodology, Y.S.; software, K.K.; validation, J.K.; formal analysis, Y.S.; investigation, K.K.; resources, K.K.; data curation, I.C.; writing—original draft preparation, Y.S.; writing—review and editing, I.C.; visualization, I.C.; supervision, I.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Filieri, R.; Raguseo, E.; Vitari, C. What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *Int. J. Hosp. Manag.* **2018**, *77*, 333–341. [[CrossRef](#)]
2. Filieri, R. What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *J. Bus. Res.* **2015**, *68*, 1261–1270. [[CrossRef](#)]
3. Huang, J.-H.; Chen, Y.-F. Herding in online product choice. *Psychol. Mark.* **2006**, *23*, 413–428. [[CrossRef](#)]
4. Zhao, X.R.; Wang, L.; Guo, X.; Law, C.H.R. The influence of online reviews to online hotel booking intentions. *Int. J. Contemp. Hosp. Manag.* **2015**, *27*, 1343–1365. [[CrossRef](#)]
5. Anagnostopoulou, S.C.; Buhalis, D.; Kountouri, I.L.; Manousakis, E.; Tsekrekos, A.E. The impact of online reputation on hotel profitability. *Int. J. Contemp. Hosp. Manag.* **2019**, *32*, 20–39. [[CrossRef](#)]
6. Phillips, P.; Barnes, S.; Zigan, K.; Schegg, R. Understanding the impact of online reviews on hotel performance: An empirical analysis. *J. Travel Res.* **2017**, *56*, 235–249. [[CrossRef](#)]
7. Yye, Q.; Law, R.; Gu, B. The impact of online user reviews on hotel room sales. *Int. J. Hosp. Manag.* **2009**, *28*, 180–182. [[CrossRef](#)]
8. Rose, M.; Blodgett, J.G. Should hotels respond to negative online reviews? *Cornell Hosp. Q.* **2016**, *57*, 396–410. [[CrossRef](#)]
9. Mudambi, S.M.; Schuff, D. Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Q.* **2010**, *34*, 185–200. [[CrossRef](#)]
10. Baek, H.; Ahn, J.; Choi, Y. Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues. *Int. J. Electron. Commer.* **2012**, *17*, 99–126. [[CrossRef](#)]
11. Huang, A.H.; Chen, K.; Yen, D.C.; Tran, T.P. A study of factors that contribute to online review helpfulness. *Comput. Hum. Behav.* **2015**, *48*, 17–27. [[CrossRef](#)]
12. Lee, S.; Choeh, J.Y. The determinants of helpfulness of online reviews. *Behav. Inf. Technol.* **2016**, *35*, 853–863. [[CrossRef](#)]
13. Hlee, S.; Lee, J.; Yang, S.-B.; Koo, C. The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *Int. J. Hosp. Manag.* **2018**, *77*, 195–206. [[CrossRef](#)]
14. Zhou, S.; Guo, B. The order effect on online review helpfulness: A social influence perspective. *Decis. Support Syst.* **2016**, *93*, 77–87. [[CrossRef](#)]
15. Chatterjee, S. Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *Int. J. Hosp. Manag.* **2020**, *85*, 102356. [[CrossRef](#)]
16. Srivastava, V.; Kalro, A.D. Enhancing the Helpfulness of Online Consumer Reviews: The Role of Latent (Content) Factors. *J. Interact. Mark.* **2019**, *48*, 33–50. [[CrossRef](#)]
17. Deutsch, M.; Gerard, H.B. A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* **1955**, *51*, 629. [[CrossRef](#)] [[PubMed](#)]
18. Lee, S.; Choeh, J. Using the Social Influence of Electronic Word-of-Mouth for Predicting Product Sales: The Moderating Effect of Review or Reviewer Helpfulness and Product Type. *Sustainability* **2020**, *12*, 7952. [[CrossRef](#)]
19. Mohammadiani, R.P.; Mohammadi, S.; Malik, Z. Understanding the relationship strengths in users' activities, review helpfulness and influence. *Comput. Hum. Behav.* **2017**, *75*, 117–129. [[CrossRef](#)]

20. Risselada, H.; de Vries, L.; Verstappen, M. The impact of social influence on the perceived helpfulness of online consumer reviews. *Eur. J. Mark.* **2018**, *52*, 619–636. [[CrossRef](#)]
21. Lankton, N.K.; McKnight, H.D. Examining two expectation disconfirmation theory models: Assimilation and asymmetry effects. *J. Assoc. Inf. Syst.* **2012**, *13*, 1. [[CrossRef](#)]
22. Oliver, R.L. A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Mark.* **1980**, *17*, 460–469. [[CrossRef](#)]
23. Burnkrant, R.E.; Cousineau, A. Informational and Normative Social Influence in Buyer Behavior. *J. Consum. Res.* **1975**, *2*, 206–215. [[CrossRef](#)]
24. Cheung, M.Y.; Luo, C.; Sia, C.L.; Chen, H. Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations. *Int. J. Electron. Commer.* **2009**, *13*, 9–38. [[CrossRef](#)]
25. Kaplan, M.F.; Miller, C.E. Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *J. Pers. Soc. Psychol.* **1987**, *53*, 306. [[CrossRef](#)]
26. Meek, S.; Wilk, V.; Lambert, C. A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *J. Bus. Res.* **2021**, *125*, 354–367. [[CrossRef](#)]
27. Lee, S.; Lee, S.; Baek, H. Does the dispersion of online review ratings affect review helpfulness? *Comput. Hum. Behav.* **2021**, *117*, 106670. [[CrossRef](#)]
28. Einar, B.; Havro, L.J.; Moen, O. An Empirical Investigation of Self-Selection Bias and Factors Influencing Review Helpfulness. *Int. J. Bus. Manag.* **2015**, *10*, 16. [[CrossRef](#)]
29. Quaschnig, S.; Pandelaere, M.; Vermeir, I. When Consistency Matters: The Effect of Valence Consistency on Review Helpfulness. *J. Comput. Commun.* **2014**, *20*, 136–152. [[CrossRef](#)]
30. Kwok, L.; Xie, K.L. Factors contributing to the helpfulness of online hotel reviews: Does manager response play a role? *Int. J. Contemp. Hosp. Manag.* **2016**, *28*, 2156–2177. [[CrossRef](#)]
31. Guo, B.; Zhou, S. What makes population perception of review helpfulness: An information processing perspective. *Electron. Commer. Res.* **2016**, *17*, 585–608. [[CrossRef](#)]
32. Ren, G.; Hong, T. Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews. *Inf. Process. Manag.* **2018**, *56*, 1425–1438. [[CrossRef](#)]
33. Henningsen, D.D.; Henningsen, M.L.M. Examining Social Influence in Information-Sharing Contexts. *Small Group Res.* **2003**, *34*, 391–412. [[CrossRef](#)]
34. Petty, R.E.; Brinol, P.; Priester, J.R. Mass Media Attitude Change: Implications of the Elaboration Likelihood Model of Persuasion. In *Media Effects*; Routledge: London, UK, 2009; pp. 141–180.
35. Xu, X.; Yao, Z. Understanding the role of argument quality in the adoption of online reviews: An empirical study integrating value-based decision and needs theory. *Online Inf. Rev.* **2015**, *39*, 885–902. [[CrossRef](#)]
36. Choi, H.S.; Leon, S. An empirical investigation of online review helpfulness: A big data perspective. *Decis. Support Syst.* **2020**, *139*, 113403. [[CrossRef](#)]
37. Kim, S.J.; Maslowska, E.; Malthouse, E.C. Understanding the effects of different review features on purchase probability. *Int. J. Advert.* **2017**, *37*, 29–53. [[CrossRef](#)]
38. Shin, S.; Du, Q.; Ma, Y.; Fan, W.; Xiang, Z. Moderating effects of rating on text and helpfulness in online hotel reviews: An analytical approach. *J. Hosp. Mark. Manag.* **2020**, *30*, 159–177. [[CrossRef](#)]
39. Wu, C.; Mai, F.; Li, X. The effect of content depth and deviation on online review helpfulness: Evidence from double-hurdle model. *Inf. Manag.* **2020**, *58*, 103408. [[CrossRef](#)]
40. Geng, S.; Niu, B.; Feng, Y.; Huang, M. Understanding the focal points and sentiment of learners in MOOC reviews: A machine learning and SC-LIWC-based approach. *Br. J. Educ. Technol.* **2020**, *51*, 1785–1803. [[CrossRef](#)]
41. Park, E.; Kang, J.; Choi, D.; Han, J. Understanding customers' hotel revisiting behaviour: A sentiment analysis of online feedback reviews. *Curr. Issues Tour.* **2018**, *23*, 605–611. [[CrossRef](#)]
42. Tian, G.; Lu, L.; McIntosh, C. What factors affect consumers' dining sentiments and their ratings: Evidence from restaurant online review data. *Food Qual. Prefer.* **2020**, *88*, 104060. [[CrossRef](#)]
43. Zeng, Z.; Zhou, Z.; Mu, X. User review helpfulness assessment based on sentiment analysis. *Electron. Libr.* **2020**, *38*, 337–351. [[CrossRef](#)]
44. Agnihotri, A.; Bhattacharya, S. Online Review Helpfulness: Role of Qualitative Factors. *Psychol. Mark.* **2016**, *33*, 1006–1017. [[CrossRef](#)]
45. Zhou, Y.; Yang, S. Roles of Review Numerical and Textual Characteristics on Review Helpfulness Across Three Different Types of Reviews. *IEEE Access* **2019**, *7*, 27769–27780. [[CrossRef](#)]
46. Mohsin, A.; Lockyer, T. Customer perceptions of service quality in luxury hotels in New Delhi, India: An exploratory study. *Int. J. Contemp. Hosp. Manag.* **2010**, *22*, 160–173. [[CrossRef](#)]
47. Presbury, R.; Fitzgerald, A.; Chapman, R. Impediments to improvements in service quality in luxury hotels. *Manag. Serv. Qual. Int. J.* **2005**, *15*, 357–373. [[CrossRef](#)]
48. Becerra, M.; Santaló, J.; Silva, R. Being better vs. being different: Differentiation, competition, and pricing strategies in the Spanish hotel industry. *Tour. Manag.* **2013**, *34*, 71–79. [[CrossRef](#)]
49. Lu, C.; Berchoux, C.; Marek, M.; Chen, B. Service quality and customer satisfaction: Qualitative research implications for luxury hotels. *Int. J. Cult. Tour. Hosp. Res.* **2015**, *9*, 168–182. [[CrossRef](#)]

50. Hlee, S. How reviewer level affects review helpfulness and reviewing behavior across hotel classifications: The case of Seoul in Korea. *Ind. Manag. Data Syst.* **2020**, *121*, 1191–1215. [[CrossRef](#)]
51. Xie, K.L.; Zhang, Z.; Zhang, Z. The business value of online consumer reviews and management response to hotel performance. *Int. J. Hosp. Manag.* **2014**, *43*, 1–12. [[CrossRef](#)]
52. Lee, M.; Lee, S.A.; Koh, Y. Multisensory experience for enhancing hotel guest experience: Empirical evidence from big data analytics. *Int. J. Contemp. Hosp. Manag.* **2019**, *31*, 4313–4337. [[CrossRef](#)]
53. Cró, S.; Martins, A.M. Hotel and hostel location in Lisbon: Looking for their determinants. *Tour. Geogr.* **2017**, *20*, 504–523. [[CrossRef](#)]
54. Long, J.S. Regression Models for Categorical and Limited Dependent Variables. *J. Am. Stat. Assoc.* **1997**, *92*, 1655.
55. Zhou, Y.; Yang, S.; Li, Y.; Chen, Y.; Yao, J.; Qazi, A. Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining. *Inf. Process. Manag.* **2019**, *57*, 102179. [[CrossRef](#)]
56. Antonakis, J.; Bendahan, S.; Jacquart, P.; Lalive, R. Causality and endogeneity: Problems and solutions. In *The Oxford Handbook of Leadership and Organizations*; Oxford Library of Psychology: Oxford, UK, 2014; Volume 1, pp. 93–117.
57. Zaefarian, G.; Kadile, V.; Henneberg, S.C.; Leischnig, A. Endogeneity bias in marketing research: Problem, causes and remedies. *Ind. Mark. Manag.* **2017**, *65*, 39–46. [[CrossRef](#)]
58. Venkatesh, V.; Thong, J.Y.; Xu, X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q.* **2012**, *36*, 157–178. [[CrossRef](#)]
59. Filieri, R.; Raguseo, E.; Vitari, C. Extremely negative ratings and online consumer review helpfulness: The moderating role of product quality signals. *J. Travel Res.* **2021**, *60*, 699–717. [[CrossRef](#)]
60. Chevalier, J.A.; Mayzlin, D. The Effect of Word of Mouth on Sales: Online Book Reviews. *J. Mark. Res.* **2006**, *43*, 345–354. [[CrossRef](#)]
61. Nakayama, M.; Wan, Y. The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. *Inf. Manag.* **2018**, *56*, 271–279. [[CrossRef](#)]