



Beyond words: unveiling the implications of blank reviews in online rating systems

Juan Pedro Mellinas¹ · Veronica Leoni^{2,3}

Received: 12 February 2024 / Revised: 24 June 2024 / Accepted: 3 September 2024
© The Author(s) 2024

Abstract

This study analyzes how review length relates to numerical scores on online platforms, conducting separate analyses for positive and negative comments and accounting for non-linearities in the relationship. Moreover, we consider the role played by blank reviews, i.e. those ratings without textual content, a topic that has been largely overlooked in previous works. Our findings suggest that blank reviews are positively correlated with higher scores, which has important implications for the ordering of reviews on online platforms. We propose that these results can be explained by social exchange theory, which suggests that less strict review policies could increase engagement and lead to a more balanced evaluation of establishments. This could offset the tendency of dissatisfied guests to disproportionately report negative experiences. Future studies should compare the composition of guest reviews on platforms adopting differing review policies.

Keywords Online reviews · Textual length · Blank reviews · Review ordering · Social exchange theory

✉ Veronica Leoni
veronica.leoni@uib.es
Juan Pedro Mellinas
mellinas@um.es

¹ Departamento de Comercialización e Investigación de Mercados, Universidad de Murcia, Murcia, Spain

² Department of Applied Economics, University of the Balearic Islands, Carretera de Valldemossa km 7.5, 07122 Palma de Mallorca, Spain

³ Center for Advanced Studies in Tourism, University of Bologna, Via Angherà, Rimini 22 47921, Italy

1 Introduction

Online reviews have become an increasingly important factor driving consumers' purchasing decisions. The impact of online generated content on consumer behaviour has been at the centre of academic interest, especially within the fields of tourism studies. Online reviews on Online Travel Agencies (OTA) websites, allow consumers to share their experiences and opinions about products and services with a wider audience. Not only do they provide valuable information to potential buyers, but they can also serve as a form of social proof, helping to shape the reputation of a company.

Most studies in the existing literature focus on the effect of numerical ratings on key aspects of products and services, such as prices (Rezvani and Rojas 2020), performance (Anagnostopoulou et al. 2019), and survivorship (Leoni 2020). Numerical scores offer consumers a simple and efficient method to evaluate the quality of products or services and make informed decisions (Acemoglu et al. 2022; Confente and Vigolo 2018), especially for experience goods whose true quality can only be determined through use (Nelson 1970). However, the accompanying text is just as important, as it allows the reviewer to provide more details and context about their experiences. Research suggests that reviews that include both ratings and text are more valuable to users than those without text (S. Lee and Choeh 2016). Additionally, longer reviews that provide more information are perceived as more helpful by the end-user (Li and Huang 2020).

While it is quite straightforward to imagine that the length of a review could affect its perceived usefulness, previous research on this topic has also explored on the relationship between length and the score magnitude. Empirical research has consistently shown that reviews containing a greater volume of text are more likely to convey negative sentiments and receive lower ratings. One plausible explanation is that negative experience tends to trigger stronger emotional responses, which in turn may motivate individuals to write more extensively about their dissatisfaction (Chevalier and Mayzlin 2006).

The design of review provisions varies significantly across online platforms. Some platforms like TripAdvisor, for tourism-related products, and IMDB, for entertainment content, have a minimum character requirement for submitting a rating. On the other hand, platforms like Amazon (for commodities), Booking.com (for tourism products), and Google (for both goods and services) provide additional options for consumers to assess products and services, allowing users to submit scores without any accompanying text. The differences in review provisions across platforms may have important implications for the quality and usefulness of the reviews. However, there is a gap in the academic literature, especially when it comes to blank reviews i.e. those rating without any textual content. Applied studies on online reviews tend to systematically exclude such reviews (Antonio et al. 2018), resulting in potentially myopic conclusions. On these premises, this examination seeks to provide a more comprehensive understanding of the workings and dynamics of online reviews, given their pivotal role in shaping consumer decision-making process.

To this purpose, we work on a rich dataset of more than 450,000 individual reviews, provided by consumers staying in hotels booked through Booking.com, the leading European OTA. Each record consists in three main parts: (i) reviewing guest characteristics; (ii) numerical score rating; (iii) (optional) written textual content. Applying linear regression model, we test the relationship between the reviews' length and the numerical score, taking advantage of the unique structure of Booking.com reviews which have separate sections for *pros* and *cons*. This allows also us to analyse positive and negative comments separately. We consider the possibility of non-linear relationships between the number of (positive and/or negative) words and the score.

Importantly, we focus on the association between numerical scores and blank reviews. This aspect underscores the originality of our work, as we reveal a significant association: blank reviews often coincide with exceptionally high scores. In this regard, we discuss the potential relationship between the increasing use of mobile devices for writing reviews and the presence of text-less the reviews. Moreover, we contribute to a critical debate on the role of review ordering, especially in the presence of a significant proportion of blank reviews, considering their tendency to correlate with high scores.

2 Literature review

Empirical research has consistently shown that the presence of text in a review enhances its value to other users (Lee and Choeh 2016; Ludwig et al. 2013). Moreover, the length of textual content has also been found to positively impact their perceived helpfulness (Li and Huang 2020; Park and Nicolau 2015). Longer reviews are not only viewed as more useful in determining product quality, but they are also perceived as more credible signal than shorter reviews (Fileri 2016).

In the realm of tourism, the motivations of users to share their experiences through online reviews have been studied using various theories from the fields of psychology, sociology, and consumer behaviour. Among them are the Social Cognitive theory (Munar and Jacobsen 2014), Extended Unified Theory of Acceptance and Use of Technology (Herrero et al. 2017), Social Identity theory (Lee et al. 2014), Social Influence theory (Oliveira et al. 2020) and Self-determination theory (Hew et al. 2017). But *Social Exchange Theory* is widely considered the most prevalent theoretical framework used to explain the motivations behind writing online reviews (Blau 1964; Homans 1958). According to this theory, individuals' decision to share information is based on a cost-benefit analysis of the sharing process (Benoit et al. 2016). Therefore, such theory has a good fit describing the propensity and motivations to share content on social media and other reviewing platforms (Cheung and Lee 2012; Wang et al. 2019; Liu et al. 2019). Within this framework, when the perceived expenses associated with sharing exceed the potential advantages, individuals are less inclined to write an online review.

The degree of consumers' satisfaction/dissatisfaction with a product or service, also plays a fundamental role. When individuals are particularly dissatisfied with a product or brand, they are more likely to express their negative experiences to

others (Bakshi et al. 2021; Eslami et al. 2018; Gretzel and Yoo 2008; Hennig-Thurau et al. 2004). This is in line with the theory of social sharing of emotions, which posits that individuals are more willing to share their negative emotions with others (Rimé 2009; Rimé et al. 1998). Consumers tend to write reviews when they have a strong desire to express themselves, and dissatisfaction appears to increase the likelihood of writing a review (Anderson 1998).

In addition to being a stronger motivator for writing a review than satisfaction, dissatisfaction may also lead to longer reviews. Researchers have studied the relationship between review length and review score for various products and services, and have found that longer reviews are often associated with lower scores in the hotel industry (Zhao et al. 2019). This pattern has also been observed in Amazon products (Eslami et al. 2018; Korfiatis et al. 2012), services reviewed on Yelp and Google (Agarwal et al. 2020; Hossain and Rahman 2022; Pashchenko et al. 2022), and other services such as insurance, financial services, and specialized drug treatment facilities (Agarwal et al. 2020; Ghasemaghaei et al. 2018; Hossain and Rahman 2022).

The widespread adoption of smartphones has significantly increased the accessibility of online review systems, which has contributed to the growth in the volume of online reviews (Burtch and Hong 2014). However, while rating on a smartphone may be more convenient and user-friendly than on a desktop computer, typing lengthy texts may not be as convenient, which can impact the quality and quantity of text-based reviews submitted through mobile devices. This is confirmed by a study using London hotel reviews, finding that the share of textless reviews is higher for smartphone than desktop devices. Also, when they contain text, it is usually less extensive and with limited analytical thinking (Park et al. 2022). This deterioration in the quality of reviews when written with mobile devices has also been identified in other studies (Mariani et al. 2019; Mudambi and Schuff 2010; Park 2018). Interestingly, Mariani et al. (2019) found a significant increase in the percentage of reviews submitted through mobile devices. At the beginning of 2015, mobile and desktop reviews were similar in percentage, but by the end of 2016, mobile reviews nearly doubled that of desktop reviews.

ReviewTrackers' recent research (2022) analyzed review length on Facebook, Google, TripAdvisor, and Yelp from 2010 to 2020. The study found that the increasing use of mobile devices is linked to a 65% decrease in review length over the decade. This trend towards mobile devices is impacting review length on prominent platforms, including Booking.com. Yoon et al.'s (2019) study on TripAdvisor data also supports these findings, indicating that reviews submitted through mobile devices tend to be shorter and of lower quality compared to those submitted through desktop computers in the tourism sector. However, the study also sought to confirm the relationship between review length and score, disaggregated by device type. Interestingly, the results showed a significant difference between desktop and mobile reviews. While mobile reviews exhibited an inverse relationship between review length and score, with shorter reviews tending to have higher scores, this was not the case for desktop reviews. In other words, there was no apparent correlation between review length and score for reviews submitted through desktop computers.

Platforms typically display reviews in a default order, which can be modified by applying different criteria, such as date and language. The length of textual comments also plays a significant role in review appearance on platforms. In the case of Google Reviews, that default order rewards reviews that contain text and photographs or that are more recent. Reviews without text or photos often rank lower in online review systems because they provide minimal information and are less helpful to potential users who are trying to make informed decisions. TripAdvisor seems to take into account solely the language and date, that is to say, in the first positions the reviews of the user's language ordered from most recent to oldest. Booking.com also has default review order criteria, which it explains on a section of its website that it first shows the reviews in the user's language and "*Anonymous reviews and those with scores but no comments will appear lower down the page*" (Booking.com 2023). This means that if a review submitted recently does not contain any text or photos, it will not be displayed on the first pages of the list of reviews that are shown by default.

The order in which reviews are presented on booking platforms can have a significant impact on consumer behavior. The appearance order can also sway an individual's opinion as prior ratings may act as a reference point. Research has demonstrated how prior ratings can impact people's perceptions and their intention to rate (Cicognani et al. 2022). The importance of considering the impact of others' opinions on individual actions is supported by both well-grounded economic and psychological theories, such as the *anchoring effect* and *social bias* (Cicognani et al. 2022; Book et al. 2016; Muchnik et al. 2013). These theories suggest that individuals are influenced by the opinions and actions of those around them, and that this influence can significantly shape their own behaviour. Blank reviews play a crucial role in this context. One might argue that if textless reviews are consistently relegated to the bottom of the list and given low scores, they may be overlooked by consumers, leading to uninformed decisions. On the other hand, if blank reviews receive high scores, their placement at the bottom of the list could potentially harm business results, as less satisfied customers' reviews will be the first to catch consumers' attention. Therefore, it is crucial to conduct a thorough analysis to determine whether blank reviews are consistently assigned by extremely satisfied or dissatisfied customers, or if they are not specifically associated with either group. This analysis will provide valuable insights into the distribution of blank reviews and their potential impact on consumer behavior.

3 Data and methodology

The following subsections provide a description of the dataset used in the analysis and the empirical strategy adopted to inspect the correlation between the review text length and the overall score.

3.1 Data

We use data collected on four European cities: Milan, Madrid, Zurich and Brussels. The data has been gathered from the Booking.com website, using a custom-designed web crawler which enables us to retrieve information about reviews both at the hotel level and at the individual level¹. From a geographical standpoint, analyzing multiple destinations improves the generality of our results and mitigate the risk of any singular destination's idiosyncratic effects influencing the results. Moreover, the set of cities considered in the current study displays a good balance between leisure and business tourism. The selection of Booking.com as platform for gathering online reviews is driven by its numerous advantageous features. Firstly, the platform utilizes a scale of 1–10 for rating, as opposed to the commonly used scale of 1–5, which allows for greater variability in the data collected (Mellinas and Martin-Fuentes 2021; Leoni and Boto-Garcia 2023). Secondly, it features two distinct sections for positive and negative feedback, providing a more accurate representation of the overall experience. This is particularly useful as current sentiment analysis software, while advanced, may not be able to accurately distinguish between positive and negative aspects of a review. Thirdly, Booking.com also provides some information on the guests' characteristics, such as the name, the type of travel party, the nationality, the length and the time of the stay. This allows for a more comprehensive analysis as it provides insights into the demographic profile of the guests, which can also be used to control for preferences and expectations of different groups of travellers. Lastly, as opposed to other platforms, Booking.com is a verified online review platform which prevent the risk of fake reviews, ensuring higher reliability and accurateness of the collected data (Figini et al. 2020). Booking.com currently requests customers to evaluate their hotel experience through an email sent after they check out, which ensure legitimacy of reviews. The evaluation process entails guests rating their overall experience on a scale of 1 to 10. Guests can also provide a non-mandatory textual comment about the positives and negatives of their stay in separate sections.

The dataset comprises approximately 458,000 individual reviews, contributed by guests who stayed in one of the 423 hotels, condo hotels, and guesthouses included in our analysis. We deliberately excluded other types of accommodations, such as private apartments, to ensure comparability in terms of critical mass of reviews. About half of the accommodation is in Milan (48%), followed by Madrid (29%), Brussels (13%) and Zurich (10%). However, in terms of individual records, Madrid has the highest number of reviews (47%), followed by Brussel (23%), Milan (21%) and Zurich (9%). Each observation has three main pieces of information: (i) travelled-related characteristics, including the guest's name (which could also be displayed as anonymous), the type of travel party, the length of stay in the accommodation, the nationality, the travel period (month and year), and the review date; (ii) an overall numeric score, on a scale from 1 to 10; (iii) a textual review composed by *pros* and *cons* sections. It is important to note that Booking.com does not required to

¹ We employed *Octoparse* software (<https://www.octoparse.com/>).

provide a textual review in order to provide a numeric evaluation of the received service. This implies that some reviews in the dataset have no textual content. Reviews are all in their original language.

As per the temporal frame covered, we considered all reviews for overnight stays in 2021 and 2022. This allows to exclude from the analysis the hotel lockdowns periods of 2020 due to Covid-19 pandemic (Leoni and Moretti 2023), and also, to some extent, allow for a higher homogeneity by avoiding the effects of changes in the Booking.com score system, implemented between late 2019 and the beginning of 2020 (Mellinas and Martin-Fuentes 2021). This approach helps us avoid the complications of mixing reviews calculated using different scales and calculation systems within the same study.

Table 1 displays the descriptive statistics of the considered sample and a brief description of the variables used in the econometric model. The overall score, which acts as our dependent variable, has been transformed into a natural logarithm to handle with its strong asymmetry, characterized by a highly left skewed distribution (Mariani and Borghi 2018). About 3,6% of reviews are left guests who decided not to reveal their identity. Such share is consistent across cities. The average percentage of reviews from domestic guests is 35%, but this distribution varies across different cities due to the different weights of domestic travel in the analyzed countries. Madrid displays the highest share of domestic travel (44%), followed by Milan and Zurich (around 23%), and Brussels (13%). On average, people stayed 2.19 days in the accommodation, again with a high degree of variation (min.1, max 39). However, the 99th percentile of the length of stay variable is less than 8 days, such distribution being common to all cities. The temporal distribution of reviews shows the classic seasonality of European destinations, with a peak of reviews during summer months. By contrast, 2022 is over represented with respect to 2021, accounting for the 65% of the sample. This is most likely linked to the fact that, during part of 2021, despite the end of lockdowns, the ongoing COVID-19 pandemic resulted in fewer guests staying in accommodations, thus leading to fewer reviews being posted. 43% of reviews belong to trips made in *couple*, followed by family (25%), solo-traveler (18%), and group (14%). We include two metrics per hotel over time: the average score and average review length (word count) up to the week before the individual review. This approach considers prior ratings and the hotel's reputation and feedback evolution in determining guests' expectations.

Table 2 presents a collection of hotel time-invariant characteristics. Based on the descriptive statistics, less than 1% of hotels are classified as one-star properties, while approximately 5% are categorized as two-star properties. The majority fall into the four-star category, accounting for 53% of the total. Furthermore, 34% of hotels are classified as four-star properties, while roughly 6% of hotels are five-star establishments. Notably, a small percentage of hotels on the Booking.com website lack a star classification. For traditional accommodations, the star rating is assigned by an official organization and transmitted to Booking.com by the hotel. In contrast, alternative accommodations are rated by the platform itself, based on various characteristics.

In terms of the type of accommodation property, hotels make up 93% of the total, followed by 7% of Condo Hotels (apartments and aparthotels). The

Table 1 Descriptive statistics of review scores and guests' profile

Variable	Description	Mean/%	Min	Max	St.Dev
<i>inscore</i>	Overall score (in natural logarithm)	2.086	0	2.303	0.307
<i>anonymous</i>	= 1 if guest is anonymous	0.036			
<i>domestic</i>	= 1 if guest is a domestic tourism	0.348			
<i>los</i>	Length of stay (in nights)	2.231	1	39	1.428
<i>month review</i>	Month of the review	7.149	1	12	3.137
<i>year review</i>	Year of the review	2021.7	2021	2022	0.476
<i>couple</i>	= 1 if travel party is <i>couple</i>	0.430			
<i>family</i>	= 1 if travel party is <i>family</i>	0.252			
<i>group</i>	= 1 if travel party is <i>group</i>	0.135			
<i>blank review</i>	= 1 if the textual review is left in blank	0.484			
<i>wordcount pros</i>	Length of the <i>pros</i> textual review (# of words)	8.026	0	384	16.020
<i>wordcount cons</i>	Length of the <i>cons</i> textual review (# of words)	2.254	0	378	10.450
<i>wordcount total</i>	Length of the overall textual review (# of words)	10.279	0	638	20.573
<i>averagelength stock</i>	Average length of the textual reviews up to the previous week (with respect to each individual record)	11.029	0	175	3.480
<i>averagescore stock</i>	Average score up to the previous week (with respect to each individual record)	2.096	0	2.303	0.085
<i>N</i>		362,988			

Table 2 Descriptive statistics of providers' characteristics

Variable	Description	Mean/%	Min	Max	St. Dev.
<i>stars</i>					
	Star rating				
Unclassified		0.018			
1 star		0.008			
2 stars		0.046			
3 stars		0.341			
4 stars		0.532			
5		0.055			
Distancekm	Distance (in km) from the city-centre (Provided by Booking.com)	0.0813	0.05	150	3.8
Parking	= 1 if property has parking	0.54			
WiFifree	= 1 if free Wi-Fi available	0.972			
Reception24h	= 1 if the property offers 24 h reception service	0.816			
Laundry	= 1 if the property offers a laundry service	0.187			
Dailycleaning	= 1 if rooms are cleaned on a daily basis	0.242			
Lift	= 1 if the property has an elevator	0.578			
Gym	= 1 if the property has a gym space	0.284			
<i>Type of property</i>					
Condo Hotel	= 1 if Condo hotel	0.074			
Guesthouse	= 1 if Guest house	0.001			
Hotel	= 1 if Hotel	0.926			
Earlycheckin	= 1 if property allows early check-in (before 2pm)	0.071			
N 362,988					

remainder of the accommodations are represented by guesthouses. Location-wise, the distance from the city center is considered for each city and ranges from 5 m to 150 km, with an average distance of 800 m.

Regarding hotel amenities, approximately 54% of hotels offer private parking, while 97% provide free Wi-Fi. Moreover, 82% of hotels offer an all-day reception service, and 18% have a laundry service. In terms of room cleaning, 24% of hotels provide daily cleaning services. Additionally, around 58% of hotels have a lift, and 28% of hotels have a gym area.

As per the key variables of the study, almost half of the reviews have no textual comment. Such share is constant over the two-year considered (No statistically significant difference in the average values for the two years under consideration ($\Pr(|T| > |t|) = 0.753$). The pros sections tend to be longer (on average 8 words) than the cons section (on average 2 words), such variables display a very pronounced standard deviations, especially for the pros section (St.Dev. 19.97). On average, reviews were one word shorter in 2022 compared to 2021, based on a simple comparison of means ($H_a: \text{diff} \neq 0$ $\Pr(|T| > |t|) = 0.0000$). Table 3 displays the percentile distribution of the review's length -related variables.

Table 3 Percentile distribution sample of the review's length

	Mean	p25	p50	p75	p90	p95
<i>wordcount pros</i>	8.01	0	1	10	24	37
<i>wordcount cons</i>	1.94	0	0	0	1	11
<i>wordcount total</i>	9.94	0	1	12	29	45
<i>N</i>	458,460					

Moreover, in Table 4, we display the average number of words for the *pros* and *cons* sections and the share of blank reviews, for different intervals of the dependent variable. We offer a descriptive analysis of the joint distribution of the overall score and review length for positive and negative aspects. This analysis provides an understanding of how review length relates to overall score. Blank reviews shed light on how often guests choose not to leave a review. This analysis can serve as a starting point for more detailed statistical analysis of the variables' correlation.

3.2 Empirical strategy

In this subsection we use the dataset described above to explore the correlation between the score evaluation and the reviews' length. For sake of clarity, it is important to underline that the analysis, and the following interpretation of results, does not allow for a causal discourse. As a matter of fact, the analytical and textual reviews are simultaneous and the correlation between them does not necessarily imply causality. On these premises, we estimate the following baseline log-linear model:

$$Y_{iht} = \alpha + \beta' X_i + \mu' W_h + \phi' ReviewStock_{ht} + \gamma TextLength_{iht} + TimeFE_t + \epsilon_{iht} \quad (1)$$

where Y_{iht} is the natural logarithm of the score left by guest i and for the hotel h at the time t (with t being the exact date of the review) X_i a vector of guests' controls, W_h is a vector of hotel time-invariant characteristics (including amenities, stars, and location), while $ReviewStock_{ht}$ account for the characteristics of the pre-existing stock of reviews (average score and text length) which varies overtime. We control also for time fixed effects. To keep an easy notation, we use a generic time t . However, it should be noted that while the time of the review is daily, the aggregation level of the review stock is weekly, while fixed effects are on a monthly basis. γ is the main parameter to be estimated, which express the correlation between the text

Table 4 Average number of words and share of blank reviews, by score level

Score	1–2	3–4	5–6	7–8	9–10
<i>word count pros</i>	12.97	8.77	7.28	6.74	8.80
<i>wordcount cons</i>	8.64	7.15	4.84	2.07	1.12
<i>% blank</i>	0.26	0.29	0.36	0.51	0.50
<i>N</i>	6594	112,010	33,747	170,547	253,909

length and the numerical score. ϵ_{ih} is a normally distributed error term. To account for auto-correlated error terms for the hotel over time, we cluster standard errors at the hotel level. The models are estimated using classic OLS regression.

In addition to the baseline specifications, we explore other aspects of the relationship between score and text length. These include:

- (i) Analyzing the length of *pros* and *cons* sections separately;
- (ii) Examining the linearity of the relationship between the number of words and score by adding quadratic terms to the regression analysis;
- (iii) Considering the association between blank reviews and the attributed score;
- (iv) Proposing an alternative model specification by excluding W_h and incorporating hotel fixed effects for better control of unobserved heterogeneity;
- (v) Exploring potential geographical heterogeneity.

4 Results

This section reports the estimates of the model in (1) as well as of the extended specifications described in the methodology section.

In Table 5, Column 1 displays a positive association between the length of the full review (*wordcount total*) and the numerical score (*inscore*). *Ceteris paribus*, additional words link to a 0.5% decrease in score. However, this measure overlooks the sentiment of the text and fails to account for the difference between blank reviews and those with a positive word count. When the analysis includes a dummy variable for text-less reviews (*blank review*), the relationship is initially positive, but reaches a turning point. This aligns with previous research indicating that dissatisfied customers tend to write longer reviews (Zhao et al. 2019). The quadratic term (*TotalWordCount Squared*) significance indicates a non-linear relationship, with a negative correlation between variables as word count increases. These findings emphasize the importance of separating blank reviews to avoid biased estimates. Figure 1 provides a visual representation of the relationship between score and word length under both specifications.

Our results reveal interesting insights regarding the association between the length of positive and negative comments and the overall score of the stay. In Column 2 we show that negative comments (*wordcount cons*) tend to be associated with lower scores, with the lowest score occurring at the 149th word. On the other hand, in Column 3, we see that lengthy positive comments (*wordcount pros*) are associated with higher scores, but this effect is also nonlinear. We observed an inverted U-shape relationship with a negative interaction parameter (*Wordcount_pros Squared*), indicating that the gain in score decreases after a certain number of words. Specifically, we identified the maximum of the study function to be at the 72th word, which we have graphically represented in Fig. 1.

In line with Lind and Mehlum (2010), we provide further evidence of a statistically significant non-linear relationship for both *pros* and *cons*. To verify this, we check two conditions: (i) the first and second derivatives have the correct sign, and

Table 5 Regression estimates

	(1)	(2)	(3)	(4)
Anonymous	-0.00932*** [-3.18]	-0.00560* [-1.88]	-0.00956*** [-3.26]	-0.00683** [-2.28]
Domestic	0.000435 [0.23]	0.0000918 [0.05]	0.0155*** [6.90]	0.0134*** [6.49]
LengthOfStay	0.00354*** [5.15]	0.00269*** [3.98]	0.00344*** [5.05]	0.00248*** [3.47]
Couple	0.0188*** [8.42]	0.0185*** [8.37]	0.0211*** [9.80]	0.0197*** [9.20]
Family	0.0280*** [10.85]	0.0286*** [11.11]	0.0298*** [11.81]	0.0282*** [11.08]
Group	0.0258*** [8.67]	0.0266*** [8.91]	0.0259*** [9.02]	0.0251*** [9.21]
1.stars	-0.00675 [-0.59]	-0.00694 [-0.59]	-0.00476 [-0.48]	-0.00351 [-0.34]
2.stars	0.0104 [1.42]	0.00974 [1.33]	0.0114 [1.54]	0.00674 [0.83]
3.stars	0.0152*** [2.67]	0.0153*** [2.71]	0.0158*** [2.60]	0.0145** [2.16]
4.stars	0.0160*** [2.72]	0.0164*** [2.80]	0.0158** [2.52]	0.0180*** [2.61]
5.stars	0.0234*** [2.96]	0.0241*** [3.06]	0.0226*** [2.77]	0.0289*** [3.32]
Distancekm	-0.000597*** [-9.47]	-0.000623*** [-9.77]	-0.000413*** [-6.31]	-0.0000568 [-0.78]
Parking	-0.00130 [-0.44]	-0.00120 [-0.41]	-0.00129 [-0.45]	0.000804 [0.27]
WiFifree	0.0103* [1.71]	0.0100* [1.76]	0.00775 [1.51]	0.00486 [1.03]
Reception24h	-0.00383 [-1.06]	-0.00305 [-0.85]	-0.00296 [-0.83]	-0.00175 [-0.50]
Laundry	0.000857 [0.22]	0.00113 [0.29]	0.000565 [0.15]	0.000355 [0.08]
Dailycleaning	0.00359 [1.10]	0.00353 [1.07]	0.00326 [1.00]	0.00327 [0.86]
Lift	-0.00185 [-0.60]	-0.00170 [-0.55]	-0.00195 [-0.65]	-0.00248 [-0.81]
Gym	0.000272 [0.07]	0.00104 [0.27]	0.00118 [0.32]	0.00151 [0.41]
Earlycheckin	0.00726 [1.60]	0.00727 [1.65]	0.00645 [1.53]	0.00407 [0.92]
Guesthouse	-0.0267*** [-3.90]	-0.0243*** [-3.68]	-0.0281*** [-4.19]	-0.0220*** [-3.05]
Hotel	-0.00997** [-2.29]	-0.00909** [-2.08]	-0.0106** [-2.40]	-0.0103** [-2.09]
TotalWordCount	0.000545*** [2.61]			

Table 5 (continued)

	(1)	(2)	(3)	(4)
TotalWordCount Squared	-0.0000107*** [-9.67]			
Blank Review	0.0560*** [14.02]	0.0914*** [23.48]	0.0342*** [10.41]	0.0627*** [16.87]
StockAverageScore	0.804*** [30.14]	0.793*** [29.37]	0.773*** [28.36]	0.718*** [25.86]
StockAverageWordCount	0.00123** [2.37]	0.000759 [1.43]	0.00168*** [3.35]	0.00149*** [2.74]
Wordcount_pros		0.00318*** [25.53]		
Wordcount_pros Squared		-0.0000222*** [-16.49]		
Wordcount_cons			-0.00535*** [-23.00]	
Wordcount_cons Squared			0.00000635*** [3.98]	
Constant	0.336*** [5.88]	0.323*** [5.57]	0.406*** [6.89]	0.496*** [8.31]
<i>N</i>	362,988	362,988	362,988	273,200
<i>POSTCODE FE</i>	YES			
<i>TIME FE</i>				
adj. <i>R</i> ²	0.075	0.075	0.091	0.068

t-statistics in brackets * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, Robust SE (clustered at the hotel level)

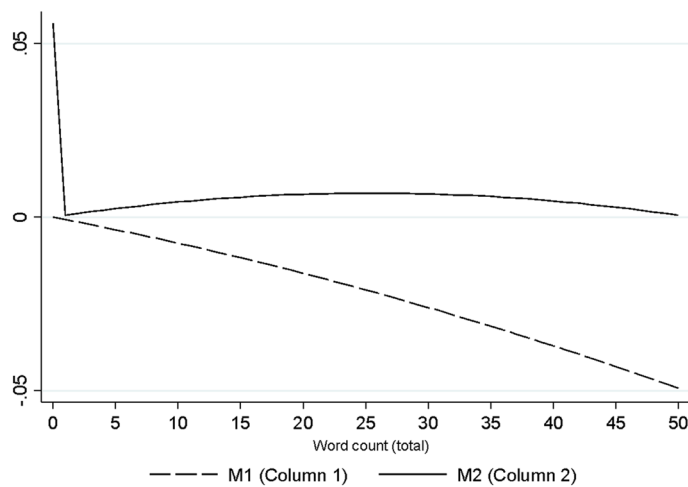


Fig. 1 Marginal effects of the model in Column 1 and Column 2 (Table 4)

(ii) the extreme points fall within the data range but not too close to the minimum and maximum values. Additionally, we calculated confidence intervals for the maximum and minimum points using the Fieller-interval method corrected for finite samples. U-test results indicate that positive comments demonstrate an inverted U-shaped relationship. However, we found insufficient evidence to suggest that negative comments have a U-shaped relationship. More specifically, the test confirms that the relationship is significant for positive comments (t-value = 14.92; 95%; Fieller interval for the extreme point: [67.246489; 76.936343]). Conversely, the minimum of the relationship between cons and the score falls outside the sample range (minimum in 421 words). These results are further confirmed by visual inspection of the two relationships, as represented in Fig. 2.

The current study explores the correlation between analytical score and blank reviews, which is a novel aspect. Results indicate that more satisfied consumers tend to provide blank reviews, all else being equal. On average, text-less reviews (*Blank Review*) display 5.76% higher scores (Marginal effect calculated using Halvorsen and Palmquist correction for log-linear model interpretation (Halvorsen and Palmquist 1980), compared with wordy reviews, regardless of the length and sentiment. Given that the great majority of reviews have a limited number of words, as previously displayed in Table 2, in Column 4 of Table 5 we have limited our analysis to the third quartile of the distribution, which includes reviews with maximum 12 words. The results show that blank reviews (*Blank Review*) still result in higher scores. To understand the impact of each word, we analyzed the effect of adding each additional word with a maximum of 12 words. We found a significant decrease in average scores as the number of words increased. On average, scores with one-word text are 17% lower compared to blank reviews, while two-word and three-word reviews are 11% and 8% lower, respectively. Figure 3 illustrates the marginal effects of word count on the score.

Overall, the analysis and corresponding graphical representation suggest that a quadratic effect alone may not adequately capture the complex relationship between the number of words and the numerical. To better represent this phenomenon, we propose a more flexible approach. Firstly, we introduce a new dummy variable to

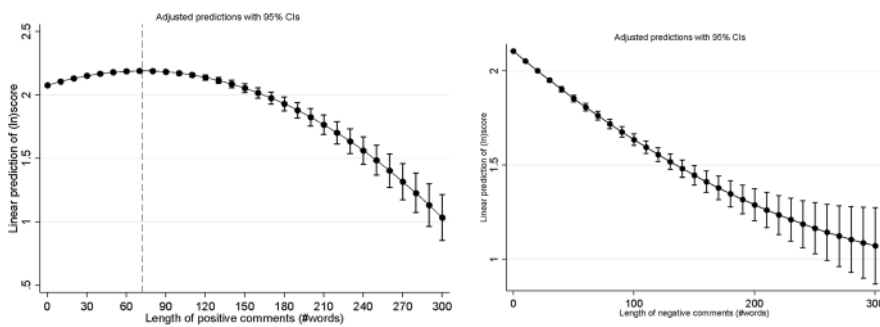


Fig. 2 Prediction margins (with 95% confidence intervals) for negative (panel a) and positive (panel b) reviews

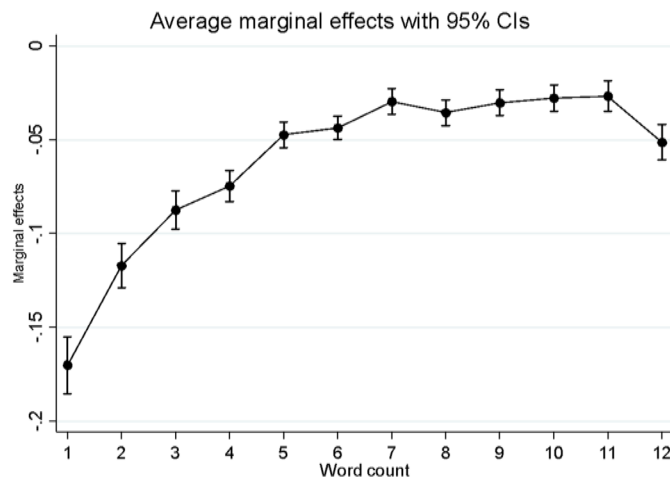


Fig. 3 Marginal effects (with 95% confidence intervals) on each additional word on the \ln score, with respect to blank reviews

distinguish blank reviews (*Blank Review*) from those with a positive count of words. Secondly, we differentiate between positive (*wordcount pros*) and negative reviews (*wordcount cons*), which significantly improves the model's performance. Our findings highlight the importance of separately considering blank reviews to enhance the explanatory and predictive power of the model.

Although not the primary focus, we provide a brief analysis of controls, including guest profiles and time-invariant hotel characteristics. *Anonymous* guests report lower scores, while guests who stay longer (*LengthOfStay*) at the hotel report higher scores. *Families* display the highest level of satisfaction, followed by *Groups* and *Couples*, compared to *Solo* travelers. *Three*, *four*, and *five-star* hotels received higher scores than unrated properties, but there was no significant difference for lower-end accommodations. Location plays a significant role, with hotels farther from the city center (*DistanceKm*) receiving lower scores. Interestingly, hotel amenities did not have a significant effect on the final evaluation, as guests were likely aware of them at the time of booking. Lastly, *Condo* tend to receive higher scores than *Hotels* and *Guesthouses*.

4.1 Robustness checks

In this section, we present a series of robustness checks that we conducted to test the validity of our results. We provide the detailed results of these tests in the Online Appendix Section.

As explained in the methodological section, we included a set of hotel characteristics in our model specification that we believe may also impact evaluations. However, we acknowledge that there may still be unobserved heterogeneity due to unknown attributes. Therefore, we conducted additional analyses, which are reported

in Table A1 (Appendix). This time, we included hotel fixed effects instead of the vector W_h . Overall, we found that our results are robust to the main specification.

The study's inference initially used clustered standard errors at the provider level but additional analyses with standard errors clustered at the postal code level showed no change in the significance level, except for the number of total words, which showed no relationship with the numerical score. Table A2 displays the results.

To conclude, we attempt to investigate the heterogeneity across various geographical regions, using data from four cities in four European countries. However, our analysis, as detailed in Table A3 of the Appendix section, did not yield any conclusive evidence of significant differences between the four countries, except for Madrid which showed a higher level of satisfaction compared to the reference category (Brussels). As an additional check, we explored whether there were significant differences in the frequency of blank reviews based on guests' nationality. In investigating how geographical location influences guest reviews, we introduced an interaction term between the 'blank reviews' variable and a dummy variable indicating whether the guest is from Europe (Europeans comprise approximately 57% of the sample). Intriguingly, results (Column5, Table A3) reveals that, on average, European guests tend to give lower scores compared to guests from other continents. However, Europeans also demonstrate a higher propensity to submit blank reviews.

5 Discussion and conclusions

In this paper we offer an empirical analysis on the relationship between the review length and the numerical score of rated posted on Booking.com platform, for hotels located in four European destinations: Milan, Madrid, Zurich and Brussels. Our findings confirm a statistically significant relationship between the number of words and the overall score rating, with longer reviews displaying average lower scores (Zhao et al. 2019; Hossain and Rahman 2022; Pashchenko et al. 2022). Besides confirming existing results, our work takes a step forward by addressing two significant gaps in the existing research on this subject.

First, we offer a deeper understanding on the above discussed relationship by analyzing positive and negative comments separately. This represents a novelty with respect with existing studies, which looked at a more generic review length. Moreover, we allow for non-linearity in the relation. We find that longer positive comments are associated with higher score, such effect displaying diminishing returns. That implies that the correlation of "more words, lower score" does not hold true if the review content is positive. However, there positive relationship between length (of *pros*) and score displays diminishing returns (maximum point at the 79th word). By contrast, longer negative comments are associated with lower score, such effect increasing for very long reviews, in accordance with what is described in previous literature.

Second, we expand our understanding of a largely neglected aspect of online user generated content: *blank reviews*. Previous studies have overlooked this area of research, systematically excluding blank reviews from empirical analysis, but our results reveal valuable insights. We not only confirm the positive correlation

between blank reviews and higher scores, which has been neglected so far, but also delve deeper into their implications. In fact, our current work precisely quantifies the differences in scores between text-with and text-less reviews. The score difference between a blank review and review with a positive number of words (regardless of the text length) is, on average, 6%. Interestingly, when restricting the analysis to short reviews, we found that a single-word review is significantly lower scored (– 17%) than a blank review.

The systematic tendency for blank reviews to display better scores has non-trivial implications. In most review platforms, blank reviews are typically positioned at the bottom of the review list, and while they still contribute to the overall score calculation, they might not carry the same weight (in terms of the value attached by prospective buyers) as reviews with text. By default, platforms display reviews with text first, while blank reviews tend to appear in the last positions. This implies that on average, the highest scores, which belong to blank reviews, are hidden. The *rationale* for ranking text-less reviews last links to their lower informative value. This approach can make sense when the goal of the platform is to provide users with useful and high-quality information. Nonetheless, our research indicates potential ramifications for providers. Specifically, our findings suggest a systematic “over visibility” of lower-scored reviews. Essentially, due to the placement of blank reviews at the bottom of the list, they receive less visibility to customers. Coupled with their correlation with higher scores, this may mean guests are primarily exposed to lower scores. This can have some sort of impact on future consumer purchasing and rating behaviors, as predicted by anchoring and social bias theories (Cicognani et al. 2022; Book et al. 2016; Muchnik et al. 2013). If users pay attention to the review scores displayed on the first page, influencing in a similar way as the overall score does, that way of ordering the reviews could have some impact on future purchase behaviors and evaluation of the users (Cicognani et al. 2022). In any case, the impact of ordering on the extent of social bias is beyond the scope of the current work. However, future studies could attempt to evaluate this effect, ideally within an experimental framework.

Another significant issue is related to the device used to provide reviews. Since the advent of smartphones, the percentage of reviews written on these devices has been gradually increasing (ReviewTrackers 2022), to the point of outweighing those written on conventional keyboard devices, especially for platforms adopting more lenient reviewing policies (Mariani et al. 2023). It is indeed true that the ease and speed of providing a review through a smartphone increases the likelihood of actually submitting one, due to the reduced opportunity cost. However, it must also be noted that writing on a smartphone may make it more challenging to draft longer and more detailed comments. There is hence a trade-off between the promptness of writing the review, which can also improve the accuracy of remembered details (Chen and Lurie 2013; Yang and Yang 2018; Healey et al. 2019), and the reduction in precision due to the use of a mobile device. In light of this, our works complement a recent work by Mariani et al. (2023) which explore the heterogeneity in the use of mobile devices in platforms displaying different review policies. The authors find that, compared to TripAdvisor, which requires a minimum number of characters, Booking.com, which allow text-free reviews, records a higher share of comments

made using mobile devices. Given that, since short and blank reviews tend to have higher ratings, which are most probably from mobile devices, the increase in smartphone usage might lead to a disproportioned increase in ratings that does not necessarily correspond to an improvement in services.

We could interpret this phenomenon under the lens of *Social Exchange* and other related theories, which explain how the motivations for sharing information also depend on the degree of effort required. In other words, the more the time (or commitment) required to write a review, the lower will be the willingness to participate in feedback sharing. Platform characterized by an easier reviewing process, will therefore register a higher number of reviews. One way to make the process easier is to allow for contributions with little or no text, as for the case of Booking.com platform. This allows certain users, who would be willing to collaborate but lack sufficient motivation to write a substantial amount of text, to also be included in the process.

The results of our study, which are consistent with those previously published, indicate that the motivation to provide highly detailed feedbacks (longer reviews) is stronger for more dissatisfied customers. Therefore, more lenient review would help counterbalancing the lower motivation of satisfied consumers. That is, given the well-known *negativity bias* in word-of-mouth (Herr et al. 1991) and the higher motivations in sharing dissatisfaction, an establishment evaluated on a platform that requires a minimum number of characters would receive a lower overall rating compared to that received on a text-free policy platform as suggested by Mariani et al. (2023).

However, the higher participation in review provision, which is attributed to lower opportunity costs, could also lead to an increase in impulsive scoring, which reflects the hotel quality with less accuracy. It is then worth considering the trade-off between having more reviews, which provides prospective buyers with more opinions, and the loss of detail associated with quickly written reviews. In other words, we believe that a higher number of reviews does not necessarily translate into higher informational value for consumers (Acemoglu et al. 2022), due to the lack of effort displayed in blank reviews. In this sense, sometimes less is more.

Our work contributes to the existing literature in four ways. First, it contributes to the extant literature on the relationship between review length and score, using a larger and more complete dataset than previous studies. Second, it prompts a rethinking of the previously established consensus regarding the correlation between review length and score by emphasising the distinction between *pros* and *cons*. Third, to the best of our knowledge, our study is the first of its kind to examine blank reviews, which, on platforms like Booking.com, make up nearly half of all reviews. Specifically, our study identifies and quantifies the difference in score associated with blank reviews. Fourth and last, offers a critical discussion on the under-explored area of the review ordering, even though is not crucial aspect of the empirical setting.

From a methodological standpoint, our work sheds light on an important aspect that should be taken into account in future studies. Specifically, our analysis and graphical representation reveal that a quadratic effect may not adequately capture the complex relationship between the number of words and numerical score. To address

this issue, we suggest relaxing this rigid structure by (i) adding a new dummy variable to the model to distinguish between blank reviews and those with a positive word count, and (ii) differentiating between positive and negative reviews. This approach significantly improves the model's performance and underscores the importance of considering blank reviews separately to improve the model's explanatory and predictive capabilities.

Our findings could be extended also to other similar websites that collect online reviews of products and services, such as Google Places or Amazon. This opens up opportunities for a critical discussion and further research about the ideal length of text that platforms should allow, both in terms of maximum and minimum character requirements, also beyond the tourism industry. Given that these decisions will be made by private entities, which currently control the entire online review ecosystem, it is likely that commercial considerations will take precedence over the goal of capturing the most accurate information possible.

Limitations of this study include the inability to differentiate between mobile and desktop review. Although the results are very similar for hotels in four cities from 4 different European countries, we cannot guarantee external validity of the results to other cities (with different city sizes and on different continents), other non-urban destinations (beach, rural, ski, etc.) or types of tourist accommodations (apartments, villas, B&B, etc.). While Booking.com was optimal, exploring other platforms with similar features is useful. A follow-up study on consumer behavior and psychology is valuable to determine if ordering, especially for blank reviews, affects consumer perceptions and rating attitudes. That follow-up study could also delve into whether the obtained results have an explanation extending beyond the realm of social exchange theory.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40558-024-00300-4>.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This study was funded by the Spanish Ministry of Science and Innovation Spanish MCIN/AEI/10.13039/501100011033/ FEDER, UE within the RevTour project [Grant Id. PID2022-138564OA-I00] "Use of online reviews for tourism intelligence and for the establishment of transparent and reliable evaluation standards".

Data availability Replication files are available upon request to the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acemoglu D, Makhdomi A, Malekian A, Ozdaglar A (2022) Learning from reviews: the selection effect and the speed of learning. *Econometrica* 90(6):2857–2899
- Agarwal AK, Wong V, Pelullo AM, Guntuku S, Polsky D, Asch DA, Muruako J, Merchant RM (2020) Online reviews of Specialized Drug Treatment facilities—identifying potential drivers of high and low patient satisfaction. *J Gen Intern Med* 35(6):1647–1653. <https://doi.org/10.1007/s11606-019-05548-9>
- Anagnostopoulou SC, Buhalis D, Kountouri IL, Manousakis EG, Tsekrekos AE (2019) The impact of online reputation on hotel profitability. *Int J Contemp Hosp Manage* 32(1):20–39
- Antonio N, de Almeida AM, Nunes L, Batista F, Ribeiro R (2018) Hotel online reviews: creating a multi-source aggregated index. *Int J Contemp Hosp Manage* 30(12):3574–3591
- Bakshi S, Gupta DR, Gupta A (2021) Online travel review posting intentions: a social exchange theory perspective. *Leisure/Loisir* 45(4):603–633. <https://doi.org/10.1080/14927713.2021.1924076>
- Benoit S, Bilstein N, Hogreve J, Sichtmann C (2016) Explaining social exchanges in information-based online communities (IBOCs). *J Service Manage* 27(4):460–480. <https://doi.org/10.1108/JOSM-09-2015-0287>
- Blau PM (1964) *Social exchange theory*. John Wiley & Sons, New York
- Book LA, Tanford S, Chen Y-S (2016) Understanding the impact of negative and positive traveler reviews: Social Influence and Price Anchoring effects. *J Travel Res* 55(8):993–1007. <https://doi.org/10.1177/0047287515606810>
- Booking.com (2023) *What order do guest reviews appear in?* Help. <https://partner.booking.com/en-gb/help/guest-reviews/general/what-order-do-guest-reviews-appear>
- Burch G, Hong Y (2014) What happens when word of mouth goes mobile? 35th International Conference on Information Systems: Building a Better World Through Information Systems, ICIS 2014. 35th International Conference on Information Systems «Building a Better World Through Information Systems», ICIS 2014. <http://www.scopus.com/inward/record.url?scp=84923435126&partnerID=8YFLogxK>
- Chen Z, Lurie NH (2013) Temporal contiguity and negativity bias in the impact of online word of mouth. *J Mark Res* 50(4):463–476
- Cheung CMK, Lee MKO (2012) What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decis Support Syst* 53(1):218–225. <https://doi.org/10.1016/j.dss.2012.01.015>
- Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: online book reviews. *J Mark Res* 43(3):345–354
- Cicognani S, Figini P, Magnani M (2022) Social influence bias in ratings: a field experiment in the hospitality sector. *Tour Econ* 28(8):2197–2218
- Confente I, Vigolo V (2018) Online travel behaviour across cohorts: The impact of social influences and attitude on hotel booking intention. *Int J Tourism Res* 20(5):660–670
- Eslami SP, Ghasemaghahi M, Hassanein K (2018) Which online reviews do consumers find most helpful? A multi-method investigation. *Decis Support Syst* 113:32–42. <https://doi.org/10.1016/j.dss.2018.06.012>
- Figini P, Vici L, Viglia G (2020) A comparison of hotel ratings between verified and non-verified online review platforms. *Int J Cult Tour Hosp Res* 2:5
- Filieri R (2016) What makes an online consumer review trustworthy? *Annals Tourism Res* 58:46–64. <https://doi.org/10.1016/j.annals.2015.12.019>
- Ghasemaghahi M, Eslami SP, Deal K, Hassanein K (2018) Reviews' length and sentiment as correlates of online reviews' ratings. *Int Res* 28(3):544–563. <https://doi.org/10.1108/IntR-12-2016-0394>
- Gretzel U, Yoo KH (2008) Use and impact of online travel reviews. *Inf Commun Technol Tour* 86:35–46
- Halvorsen R, Palmquist R (1980) The interpretation of dummy variables in semilogarithmic equations. *Am Econ Rev* 70(3):474–475
- Healey MK, Long NM, Kahana MJ (2019) Contiguity in episodic memory. *Psychon Bull Rev* 26(3):699–720
- Hennig-Thurau T, Gwinner KP, Walsh G, Gremler DD (2004) Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *J Interact Mark*, 18(1), Article 1.

- Herr PM, Kardes FR, Kim J (1991) Effects of word-of-mouth and product-attribute information on persuasion: an accessibility-diagnostics perspective. *J Consum Res* 17(4):454–462
- Herrero Á, San Martín H (2017) Explaining the adoption of social networks sites for sharing user-generated content: a revision of the UTAUT2. *Comput Hum Behav* 71:209–217
- Hew TS, Syed A, Kadir SL (2017) Applying channel expansion and self-determination theory in predicting use behaviour of cloud-based VLE. *Behav Inform Technol* 36(9):875–896
- Homans GC (1958) Social Behavior as Exchange. *Am J Sociol* 63(6):597–606. <https://doi.org/10.1086/222355>
- Hossain MS, Rahman MF (2022) Detection of potential customers' empathy behavior towards customers' reviews. *J Retailing Consumer Serv* 65:102881. <https://doi.org/10.1016/j.jretconser.2021.102881>
- Korfiatis N, García-Bariocanal E, Sánchez-Alonso S (2012) Evaluating content quality and helpfulness of online product reviews: the interplay of review helpfulness vs. review content. *Electron Commer Res Appl* 11(3):205–217. <https://doi.org/10.1016/j.elerap.2011.10.003>
- Lee S, Choeh JY (2016) The determinants of helpfulness of online reviews. *Behav Inform Technol* 35(10):853–863. <https://doi.org/10.1080/0144929X.2016.1173099>
- Lee H, Reid E, Kim WG (2014) Understanding knowledge sharing in online travel communities: antecedents and the moderating effects of interaction modes. *J Hospitality Tourism Res* 38(2):222–242
- Leoni V (2020) Stars vs lemons. Survival analysis of peer-to peer marketplaces: the case of Airbnb. *Tour Manag* 79:104091
- Leoni V, Boto-García D (2023) Apparent and actual hotel scores under Booking. Com new reviewing system. *Int J Hospitality Manage* 111:103493
- Leoni V, Moretti A (2023) Customer satisfaction during COVID-19 phases: the case of the venetian hospitality system. *Curr Issues Tourism*, 1–17
- Li M, Huang P (2020) Assessing the product review helpfulness: affective-cognitive evaluation and the moderating effect of feedback mechanism. *Inf Manag* 57(7):103359. <https://doi.org/10.1016/j.im.2020.103359>
- Lind JT, Mehlum H (2010) With or without U? The appropriate test for a U-Shaped Relationship*. *Oxf Bull Econ Stat* 72(1):109–118. <https://doi.org/10.1111/j.1468-0084.2009.00569.x>
- Lind J, Mehlum H (2019) *UTEST: Stata module to test for a U-shaped relationship*
- Liu X, Zhang Z, Law R, Zhang Z (2019) Posting reviews on OTAs: motives, rewards and effort. *Tour Manag* 70:230–237. <https://doi.org/10.1016/j.tourman.2018.08.013>
- Ludwig S, de Ruyter K, Friedman M, Brüggem EC, Wetzels M, Pfann G (2013) More than words: the influence of affective content and linguistic style matches in Online Reviews on Conversion Rates. *J Mark* 77(1):87–103. <https://doi.org/10.1509/jm.11.0560>
- Mariani MM, Borghi M (2018) Effects of the Booking. Com rating system: bringing hotel class into the picture. *Tour Manag* 66:47–52
- Mariani MM, Borghi M, Gretzel U (2019) Online reviews: differences by submission device. *Tour Manag* 70:295–298. <https://doi.org/10.1016/j.tourman.2018.08.022>
- Mariani MM, Borghi M, Laker B (2023) Do submission devices influence online review ratings differently across different types of platforms? A big data analysis. *Technol Forecast Soc Chang* 189:122296
- Mellinas JP, Martin-Fuentes E (2021) Effects of Booking. Com's new scoring system. *Tour Manag* 85:104280
- Muchnik L, Aral S, Taylor SJ (2013) Social Influence Bias: a randomized experiment. *Science* 341(6146):647–651. <https://doi.org/10.1126/science.1240466>
- Mudambi SM, Schuff D (2010) Research note: what makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Q* 34(1):185–200. <https://doi.org/10.2307/20721420>
- Munar AM, Jacobsen JKS (2014) Motivations for sharing tourism experiences through social media. *Tour Manag* 43:46–54
- Nelson P (1970) Information and consumer behavior. *J Polit Econ* 78(2):311–329
- Oliveira T, Araujo B, Tam C (2020) Why do people share their travel experiences on social media? *Tour Manag* 78:104041
- Park Y-J (2018) Predicting the helpfulness of online customer reviews across different product types. *Sustainability* 10(6):Art6. <https://doi.org/10.3390/su10061735>
- Park S, Nicolau JL (2015) Asymmetric effects of online consumer reviews. *Annals Tourism Res* 50(Supplement C):ArtSupplementC. <https://doi.org/10.1016/j.annals.2014.10.007>
- Park K, Kim H-J, Kim JM (2022) The effect of mobile device usage on creating text reviews. *Asia Pac J Mark Logistics ahead-of-print(ahead-of-print)*. <https://doi.org/10.1108/APJML-11-2021-0838>

- Pashchenko Y, Rahman MF, Hossain MS, Uddin MK, Islam T (2022) Emotional and the normative aspects of customers' reviews. *J Retailing Consumer Serv* 68:103011. <https://doi.org/10.1016/j.jretconser.2022.103011>
- ReviewTrackers (2022) *2022 Online Reviews Statistics and Trends: A Report by ReviewTrackers*. ReviewTrackers. <https://www.reviewtrackers.com/reports/online-reviews-survey/>
- Rezvani E, Rojas C (2020) Spatial price competition in the Manhattan hotel market: the role of location, quality, and online reputation. *Manag Decis Econ* 41(1):49–63
- Rimé B (2009) Emotion elicits the social sharing of emotion: theory and empirical review. *Emot Rev* 1(1):60–85. <https://doi.org/10.1177/1754073908097189>
- Rimé B, Finkenauer C, Luminet O, Zech E, Philippot P (1998) Social sharing of emotion: new evidence and new questions. *Eur Rev Social Psychol* 9(1):145–189. <https://doi.org/10.1080/14792779843000072>
- W Anderson E (1998) Customer satisfaction and Word of Mouth. *J Service Res* 1(1):Art1. <https://doi.org/10.1177/109467059800100102>
- Wang Y, Xiang D, Yang Z, Ma S, Sara (2019) Unraveling customer sustainable consumption behaviors in sharing economy: a socio-economic approach based on social exchange theory. *J Clean Prod* 208:869–879. <https://doi.org/10.1016/j.jclepro.2018.10.139>
- Yang Y, Wu L, Yang W (2018) Does time dull the pain? The impact of temporal contiguity on review extremity in the hotel context. *Int J Hospitality Manage* 75:119–130
- Yoon Y, Kim AJ, Kim J, Choi J (2019) The effects of eWOM characteristics on consumer ratings: evidence from TripAdvisor.com. *Int J Advertising* 38(5):684–703. <https://doi.org/10.1080/02650487.2018.1541391>
- Zhao Y, Xu X, Wang M (2019) Predicting overall customer satisfaction: big data evidence from hotel online textual reviews. *Int J Hospitality Manag* 76:111–121

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.