Online Reviews and Reading Strategies: The Airbnb Case.

Abstract

The platform economy mostly relies on implementing a trust generating system, through rating and reputation mechanisms. Our research aims to arrive at a better understanding of the dynamics of online bookings on the Airbnb platform through the lens of online reviews. The dataset covers the available offers in Paris in June 2019 and includes all the reviews written in French (282,057 reviews). The final dataset includes 30 variables and 31,090 offers. Several nested linear regressions are compared, which include the characteristics of the offer, host strategy, the signals regarding host quality, ratings and sentiment analysis. The results confirm an interaction effect between reviews expressing a positive sentiment and reviews expressing a negative sentiment: offers with mixed content generate the highest booking rate.

Key words: platform economy, online reviews, signal theory, sentiment analysis, reading strategies

> Sophie BALECH
University Lecturer and Researcher,
IAE Amiens – CRIISEA (EA 4286),
Pôle Universitaire Cathédrale BP 2716
80027 Amiens cedex 1,
+ 33 6 73 98 63 78,
sophie.balech@gmail.com

> Christophe BENAVENT
University Professor,
Université Paris Dauphine - PSL – CEROS (EA 4429),
200 avenue de la République
92001 Nanterre Cedex,
+ 33 6 23 23 62 56,
christophe.benavent@gmail.com

> Pauline de PECHPEYROU
University Professor,
Université Paris-Est - IRG (EA 2354),
Place de la porte des champs - Route de Choisy
94010 Créteil Cedex,
+33 6 24 42 46 71,
pauline.de-pechpeyrou@u-pec.fr

Introduction

Platform economies, such as Uber and Airbnb, are largely based on a trust generating mechanism, i.e. the establishment of a system for generating trust in relationships, with rating and reputation mechanisms (Zhou, Dresner, and Windle, 2008). These platforms have the specificity of a dual evaluation system, by suppliers and customers, which can lead to strategic behavior that can bias the quality of the signal (Masclet and Pénard, 2012). Because of their major impact on decision making and business performance, online reviews attract a great deal of interest, both at a theoretical and managerial level.

At a theoretical level, the availability of aggregate data (cinema audience statistics, online sales rankings) coupled with the availability of average indicators on the reviews expressed, initially oriented research towards sectors such as the cinema (Larceneux, 2007), pure players (Chevalier and Mayzlin, 2006) and online auction sites (Cabral and Hortacsu, 2010). This research highlights the joint impact of the volume of reviews and their valence on online sales (Chevalier and Mayzlin, 2006; Ren et al. 2018). From a managerial point of view, the crucial importance of online reviews explains the interest that customer relations managers have in them. While they are primarily sources of immediate knowledge, managers are also sensitive to the problem of “false reviews” and ways to counter or eliminate them. Several approaches have been proposed to identify false reviews (Hu et al., 2012; Munzel, 2015), including the label awarded by a third party source (Consumer Reports). Similarly, there is a strong temptation to respond even if the strategies are not obvious: in the hotel sector, it is relevant to respond to a negative review, while responding to a positive review is similar to a promotional strategy, generating reaction among internet users (Wang and Chaudhry, 2018).

Much of the now extensive research on customer reviews has focused on assessing the effect of evaluations (both quantitative and textual) on consumer response. Causality tests have established a two-way relationship between the volume of negative reviews and the rank of the item on a commercial web site, with only a marginal relationship for positive reviews (Ren et al., 2018). Thus, it is the negative reviews that play a predominant role because they are attributed to the experience itself and not to the personality of the reviewer (Chen and Lurie, 2013).

On the other hand, little research has focused on the reading strategies that underlie the internet user’s decision making. Qualitative research has shown that the motivations for consulting reviews online are essentially utilitarian since these reviews are primarily used to finalize a booking session (Séré de Lanauze and Siadou-Martin, 2018). Two levels of information are taken into account: on the one hand, the corpus of comments considered as a whole, qualified in terms of size (number of comments), trend valence (rather positive or negative) and consistency, and on the other hand, the information elements specific to each message taken individually (source, valence, length and content in terms of substance and form). Following on from this work, our research aims to test on real data the effects of the comments submitted concerning the attractiveness of an offer. As the internet user is not able to process all the available information, a sample-based reading strategy is employed. As a few samples of reviews are read, the information accumulates and it is assumed, in the spirit of Thaler’s (1985) mental accounting, that the final evaluation results from a balance between negative and positive feelings. Extremely positive reviews may call into question their veracity (too good to be true) (Maslowska, Malthouse and Bernritter, 2017) while extremely negative reviews discourage the internet user from making a reservation. We are therefore seeking to test the hypothesis of an interaction effect between positive and negative reviews using real data extracted from the Airbnb platform. All the offers available in Paris in June 2019 were collected as well as all the reviews in French. Our research therefore has three specific objectives:

1. To underline the interest of taking into account, beyond the aggregate quantitative elements (volume and valence of the ratings), the analysis of the sentiment expressed.
2. To establish the effects of mitigation between positive and negative reviews.
3. To highlight differentiated reading strategies based on situational variables such as the time horizon.

For this, different nested regression models will be compared, taking into account: (1) housing characteristics, (2) host offering strategy, (3) host related quality signals, (4) online reviews, and (5) sentiment analysis. The results support the proposed theory regarding online review reading strategies: offers with “mixed” content in customer reviews lead to the highest booking rates.

Theoretical framework

How does an internet user evaluate the attractiveness of the offers available on CtoC platforms? Which clues does he/she rely on to assess the quality of properties offered for rent? How does a more in-depth reading of the reviews submitted support or contradict an initial assessment? In addition to the aggregate indicators indicating the quality of a property, the reviews left by internet users, and in particular the feelings they express, can contribute to the judgment formed by the internet user of the attractiveness of an offer.

Trust signals

Research on online reviews highlights the preponderant role of social influence, beyond any judgment that the receiver may make on his or her own based on the product description (Tran, 2015). In a first approach, the internet user in search of goods or services can rely on several trust "signals" relating to the goods themselves or to the person offering them.

Initial research conducted on the cinema and online auction site sectors has established the role of two key indicators in explaining the level of sales: the volume of reviews issued and the average score given. A study of the reviews on the Allociné site shows that after the launch week, buzz spreads amongst internet users: the number of internet users in the first week is a significant predictor of attendance after the first week (Larceneux, 2007). Similarly, the average score is positively correlated to sales on pure player sites (Chevalier and Mayzlin, 2006). A closer examination of the effects of the scores awarded highlights the importance of negative reviews ("one star"), due to their greater scarcity (Ren et al., 2018). The first negative review received by a seller is particularly crucial since it influences the rate of sales growth, which becomes negative (Cabral and Hortacsu, 2010). Beyond the signals that testify to the quality of goods, there are trust labels that make it possible to identify sellers within a collaborative platform. For example, the "Superhost" status on Airbnb rewards the most experienced and highest rated hosts on Airbnb.

These trust signals allow an initial selection to be made by the internet user, which can be coupled with a more in-depth analysis of the offers thus selected after filtering. Indeed, recent research suggests going beyond indicators of the volume and valence of reviews to also take into account the length of the review (Fink et al., 2018) or the sentiment expressed (Wang et al., 2018). Work of a more qualitative nature confirms the consideration of these two levels of information, with the internet user relying both on aggregate indicators and on elements of information specific to each message taken individually (Séré de Lanauze and Siadou-Martin, 2018).

The contribution of sentiment analysis

The traces left on the web, whether in the form of written reviews or simple emoticons reflecting the writer's state of mind, have the advantage of being abundant, independent and spontaneous but unstructured (Moscarola and Boughzala, 2016). To analyze these corpora, it seems indispensable to mix traditional methods (based on lexical properties) with ad hoc semantic approaches (construction and application of thesauri) and sentiment analysis.

Sentiment analysis can be approached in two ways: a machine-learning approach and a lexicon-based approach. The first one focuses on the identification of sentences or proposals with evaluative expressions and uses repertoires of positive or negative terms. The discursive or summative composition (Chardon, 2013) of these elements makes it possible to establish the more or less positive or negative character of a review. This approach generally involves two stages (Alkalbani et al., 2016): a pre-processing stage for raw reviews, followed by a stage in which the reviews are classified according to their polarity. To this end, a small sample is used to train the SVM (Support Vector Machine) classification algorithm to measure the accuracy of the prediction. Based on a mapping of emojis according to the sentiment expressed (Novak et al., 2015), Crépin and Ngobo developed a neural model capable of learning the vocabulary associated with the presence of positive or negative emojis (expressing a sentiment). The second approach, based on a lexicon of pre-evaluated words or expressions, makes it possible to determine the feelings contained in texts by simply counting occurrences of these words or expressions (Mohammad and Turney, 2013; Pennebaker et al., 2015). While sentiment analysis has been used primarily for descriptive purposes, it has also been used to explain product sales (Li et al., 2019), to detect false reviews (Hu et al., 2012), and to identify the relative strengths and weaknesses of a product as expressed through consumer reviews (Wang et al., 2018).

Past research has established the importance of negative information in the internet user's decision-making process (Chevalier and Mayzlin, 2006). An attribution mechanism is at work to explain the negative relationship between the valence of the review and its usefulness: positive reviews are attributed more to the personality of the reviewer while negative reviews are attributed more to experience with the product or service (Chen and Lurie, 2013). Compared to "classic" commercial web sites, CtoC platforms are characterized by a greater empathy felt between the provider and the customer (Pera et al., 2019), which reduces the propensity to write a negative review.

Both positive and negative topics have an effect on sales that is mediated by the rating given to the product or service (Li et al., 2019). The presence of emotions in the reviews submitted also influences ratings through an attribution mechanism (Kim and Gupta, 2012): a negative review with a high emotional content will be attributed to the irrationality of its writer, unless several negative reviews express similar emotions. An offer that receives several negative reviews with strong emotional content will be rejected, since the convergence of emotions increases the informational value of the reviews and therefore, if they are negative, the feeling of making a bad choice. While positive comments are obviously desirable, one can however imagine that exclusively positive comments can lead to mistrust. The attribution mechanism discussed above (Chen and Lurie, 2013) can relate highly enthusiastic comments to the personality or intent of the writer. On the other hand, the absence of negative emotions expressed in the comments may create doubt about the credibility of the reviews.

The role of reviews in the reading and decision-making process

The now abundant research on customer reviews has focused mainly on assessing the effect of evaluations (quantitative and textual) on consumer response, but few studies have considered the reading strategies that lead to decision making. However, Afflerbac and Cho (2009) provide a useful synthesis, showing that reading is not limited to the comprehension of words and more complex phrases, but also includes constructive strategies where the choice of texts is critical. Reading is also a decision-making process that leads to filtering and selecting the elements that will receive more attention.

From the repeated observation that positive and negative reviews have an asymmetrical effect, we draw the idea that their evaluation is subject to two independent processes that proceed by elimination in the tradition of attribute elimination models (Laurent, 2007). In our case, the attributes are the positive and negative feelings perceived in the skim reading of texts, at least the most salient ones. The rule used would thus be to penalize offers whose reviews are frequently very negative, because they confirm doubt, but also, less intuitively, those whose reviews are frequently very positive, because they may also be a sign of misleading information (too good to be true). Research based on actual sales data establish that, contrary to popular belief, higher scores do not always lead to higher sales (Maslowska, Malthouse and Bernritter, 2017). On the commercial web sites studied, the probability of purchase increases with an average rating up to 4.2-4.5 stars, and then decreases (especially when the average rating is close to 5 stars). We thus formulate the hypothesis that the presence of exclusively positive or negative reviews leads to fewer reservations of the property.

In a more formal way:

H1a: The density of very negative content in the reviews has a negative impact on the degree of attractiveness of an offer on a C2C platform.

H1b: The density of very positive content in the reviews has a negative impact on the degree of attractiveness of an offer on a C2C platform.

However, in this sample-based reading strategy, information is accumulated and it is assumed, in the spirit of Thaler's (1985) mental accounting, that the final evaluation is the result of a balance between negative and positive feelings. The value of the estimated density of very positive or negative content most certainly depends on their interaction. If there are only positives or negatives, then the evaluation is bad, and the option is discarded. If there are neither, then the offer presents less risk but little appeal. If, on the contrary, the density of negatives and positives is simultaneously high, the attractiveness of the offer is more pronounced and will result in a higher probability of choice. We therefore formulate the hypothesis of an interaction effect between the density of very positive and very negative sentiments, according to which a property offered on a CtoC platform will be all the more attractive if it has been the subject of reviews with "mixed" content, with positive polarity overall.

H2: The concomitant presence of positive and negative sentiments has a positive impact on the degree of attractiveness of an offer on a C2C platform

The work of Lallement (2010) confirms the principle of selection: in a time pressure situation, the number of items of information consulted decreases. As expected, time pressure leads to a consumer selection of the number of items of information consulted but does not deprive the price attribute of its dominant role (Lallement and Zollinger, 2013). We can therefore hypothesize that when the purchase is made at a distance, the internet user will take more information into account and the role of emotional content will be increased.

H3: The role of emotional content increases with the degree of planning in the choice of an offer on a C2C platform.

Methodology

Data

The data processed comes from insideairbnb.com, which contains information about the accommodation offered for rent in more than 50 cities and user reviews associated with these rental ads. These data sets are produced by scraping airbnb.com. The scope of the study focused on the city of Paris, and the data was scraped on 6 June 2019. They include all the information relating to the listings active on that date and visible to consumers, in particular, the description of the properties offered for rent, prices, dates of availability of the offers, but also the average rating of the listing, the number of reviews received and the entire corpus of reviews written (more than one million reviews, of which 60% were written in English, 30% in French and 15% in other languages). These data are divided into three sets that were used to build the analysis models:

• listings, which includes all the variables related to an offer and its owner;
• reviews, which includes all the reviews associated with the listings;
• calendar, which includes the dates on which each listing is open for booking or not.

Variables construction

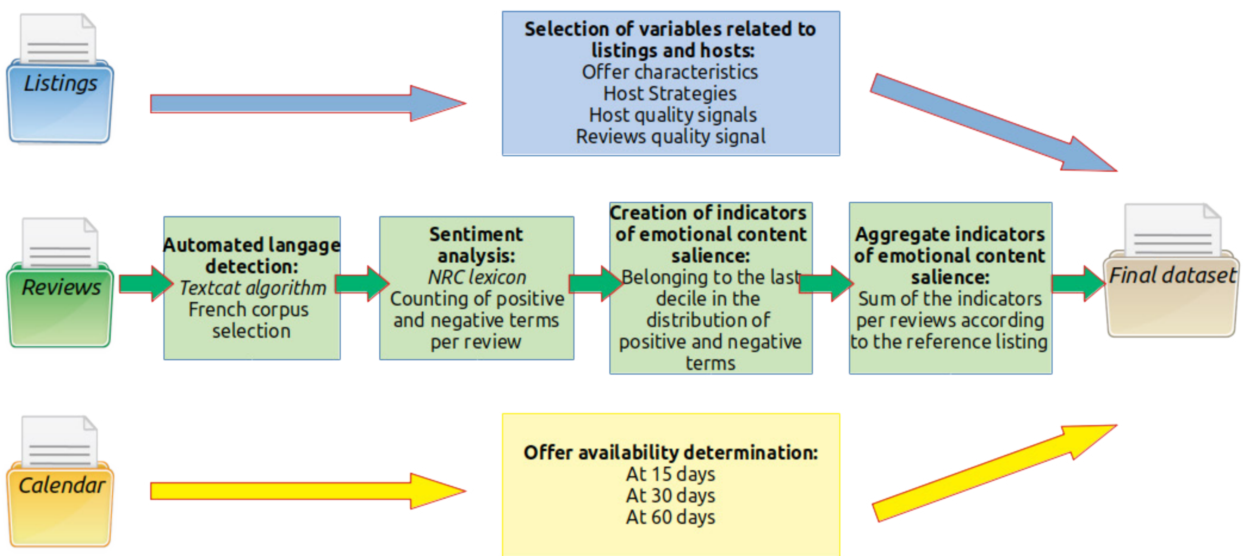
The final data set was constructed in three stages, shown in Figure 1. First, we extracted the variables related to the listings and their owner from the listings file by grouping them into four groups: the characteristics of the accommodation (neighborhood, type of rental, number of beds); the host's strategy (price, minimum number of nights, presence of house rules, deposit, household and cancellation fees, possibility of instant booking); the quality signals related to the host (total number of listings, Superhost status, identity verification by the platform, photo, length of time on the platform) and the quality signals related to the reviews (number of reviews, average score).

Then, we used the review file and text analysis methods to create indicators of the emotional content of the reviews. We first extracted the corpus of French reviews from the initial file using an n-gram-based categorization tool, the textcat algorithm (Hornik et al., 2013) available under R. For each listing, a maximum of 50 reviews (almost complete French corpus, corpus of 267,444 reviews out of 282,057 French reviews) were selected and processed using text analysis methods performed with R to determine their emotional content. More specifically, a dictionary approach was used using the NRC dictionary (Mohammad and Turney, 2013). Each term was annotated according to the feeling it expressed (positive, negative or neutral), making it possible to calculate a positive and negative feeling score for each review by counting. In order to take into account the skim type reading of online reviews by internet users when making a choice, we were interested in particularly expressive reviews, the content of which belonged to 10% of the most positive or negative reviews (see Appendix A: distribution of positive and negative scores in the corpus). These reviews act as salient elements, easily attracting attention and more strongly affecting the choice.

Finally, the calendar file made it possible to calculate the availability rate of an offer. The attractiveness of an Airbnb offer for the consumer was measured through the number of days the offer was available for booking, 15 days ahead (low degree of planning), 30 days ahead (medium degree of planning) and 60 days ahead (high degree of planning). The greater this number of days, the less the offer is reserved, and the lower its degree of attractiveness. Conversely, an offer with little or no availability is assumed to be more attractive to consumers.
Figure 1 shows the process of building up the final dataset. In total, it includes 31,090 Airbnb listings located in Paris and in operation on 6 June 2019, and 23 variables.

Figure 1

Data processing operations to build up the final dataset



Models

To explain the availability rate of the Airbnb offers, we used 5 nested linear models. The variables were added in 5 different blocks:

1. The specific characteristics of the offer (neighborhood, type of accommodation – entire property, single room, shared room –, number of beds),
2. The host's offer strategy (the minimum number of nights for a reservation, the price per night, the presence or not of house rules, the presence or not of a deposit, the presence or not of additional cleaning costs, the possibility of booking the accommodation instantaneously, the cancellation conditions – three levels: strict, moderate or unconditional),
3. Host related quality signals (total number of host listings, whether the host's identity has been verified by the platform, presence or not of a profile picture, recognition of the platform (whether the host has Superhost status or not), length of time of the host on the platform),
4. Quality signals from the reviews (number of reviews, average rating of the listing, time elapsed since the last review – indicator of recent activity),
5. Raw signals related to the reviews expressed by the salience of their emotional content (number of positive reviews belonging to the 10% most positive reviews, number of negative reviews belonging to the 10% most negative reviews).

Three time horizons were taken into account to construct the dependent variables: availability at 15 days, reflecting rapid planning, at 30 days, reflecting average planning, and availability at 60 days, reflecting long-term planning.

A total of five linear regression models were compared (Table 1) on three availability dependent variables: at 15 days, 30 days and 60 days. Fifteen regressions were performed.

Table 1

Linear regression models

Model	Expression	Variables *(name given in models)*
Model 1 lm1	avail. {15d; 30d; 60d} = Variables related to accommodation characteristics	<u>The accommodation characteristics:</u> - the number of beds available *(bed)* - the type of rental (whole apartment, room with the inhabitant, shared room) *(room)* - the neighborhood of the listing *(quartier)*
Model 2 lm2	avail. {15d; 30d; 60d} = Model 1 + host strategy variables	<u>Host strategy variables:</u> - the minimum number of nights for a reservation *(minimum_nights)* - the price of the night *(prix)* - the presence or absence of rules of procedure *(regl)* - the presence or absence of a deposit *(caution)* - whether or not there are additional cleaning costs *(menage)* - the possibility to book the accommodation instantly *(res_inst)* - cancellation conditions (strict, moderate or unconditional) *(annul)*

Model 3 lm3	avail. {15d; 30d; 60d} = Model 2 + host characteristics variables	Host quality signals: - host total listings count *(host_total_listings_count)* - if the identity has been verified by the platform *(check _identity)* - the presence or not of a profile picture *(profile_pic)* - recognition of the platform (the host has Superhost status or not) *(superhost)* - the host's length of time on the platform *(anciennete)*
Model 4 lm4	avail. {15d; 30d; 60d} = Model 3 + review related variables	Aggregated signal related to the reviews: - the average rating of the listing *(review_scores_rating)* - the number of reviews *(number_of_reviews)* - the date of the last review (indication of a recent activity) *(dernier_com)*
Model 5 lm5	avail. {15d; 30d; 60d} = Model 4 + sentiment related variables	Raw signals related to the reviews: - number of positive reviews belonging to the top 10% of the most positive reviews *(top_pos50)* - number of negative reviews belonging to the top 10% of the most negative reviews *(top_neg_50)* - interaction effect between the two indicators

Results

Model fit comparison

Table 2 compares the main performance indicators of the models tested.

- The first observation is that the longer the prediction horizon, expressed in days, the better the prediction quality. This can be explained by the more planned nature of visitor behavior. In the short term, reservations are made on a choice more constrained by circumstances. The weight of the sentiment of the reviews would be all the greater as the decision is subject to more thorough and prepared deliberation.
- The second finding is the improvement in fit as blocks of variables are added. The introduction of variables related to the host's strategy greatly improves the prediction quality of the model (lm2 model), which is explained by the major role of price in the decision process. The introduction of features related to online reviews (lm4 model) significantly improves the fit, as does the addition of sentiments (lm5 model).
- The quality of the model improves with the addition of the sentiment indicators of the reviews and with the availability period selected. These initial results confirm the importance of social influence in the choice of a purchase whose quality is difficult to assess a priori.

Table 2

Comparison of fit performance of linear regression models

	df	AIC	BIC	R^2 Adj.	RMSE
			15-day		
	df	**AIC**	**BIC**	**R^2 Adj.**	**RMSE**
lm1	28	87 354,12	87 587,77	0,029	0,986
lm2	36	84 949,04	85 249,45	0,101	0,948
lm3	41	84 845,97	85 188,10	0,104	0,946
lm4	44	84 272,94	84 640,10	0,121	0,938
lm5	47	84 156,33	84 548,53	0,124	0,936
			30-day		
	df	**AIC**	**BIC**	**R^2 Adj.**	**RMSE**
lm1	28	87 316,59	87 550,24	0,030	0,985
lm2	36	84 757,43	85 057,84	0,107	0,945
lm3	41	84 651,22	84 993,35	0,110	0,943
lm4	44	84 014,44	84 381,61	0,128	0,934
lm5	47	83 874,42	84 266,61	0,132	0,932
			60-day		
	df	**AIC**	**BIC**	**R^2 Adj.**	**RMSE**
lm1	28	87 146,99	87 380,64	0,035	0,982
lm2	36	84 064,79	84 365,20	0,126	0,935
lm3	43	83 961,45	84 303,58	0,129	0,933
lm4	45	83 133,84	83 501,00	0,152	0,921
lm5	47	82 983,94	83 376,13	0,157	0,918

Examination of the parameters

Table 3 shows the results of the estimates of the standardized coefficients of the 5 models for a 60-day availability. The average rating of the listing has a negative effect (lm4: β = -0.066, p-value = 0.000; lm5: β = -0.060, p-value = 0.000) on availability, which seems logical, a positive rating sends a quality signal to consumers who will therefore tend to turn to these listings for their stay. However, availability increases with the number of reviews of a listing (lm4: β = 0.150, p-value = 0.000; lm5: β = 0.106, p-value = 0.000), which probably reflects the active strategy of the host who gets more reviews the more he/she rents his/her accommodation and therefore it is offered more for booking.

Role of sentiment indicators

The coefficient of the negativity indicator is positive and significant (15-day availability: β = 0.073, p-value = 0.008; 30-day availability: β = 0.079, p-value = 0.008; 60-day availability: β = 0.079, p-value = 0.008). The presence of several converging negative reviews has a negative impact on the attractiveness of the offers, so **we validate hypothesis H1a.**

The coefficient of the positivity indicator is positive and significant (15-day availability: β = 0.033, p-value = 0.007; 30-day availability: β = 0.037, p-value = 0.007; 60-day availability: β = 0.041, p-value = 0.007). The positive sign of this coefficient can be explained by the low credibility associated by consumers with reviews that are too positive, perceived as not being genuine or sincere ("too good to be true") and/or by a very strong commitment of the host to the platform (high availability, high quality of stay). **Hypothesis H1b is therefore validated.**

Table 3

Comparison of standardized results of regression models on 60-day availability.

	lm1		lm2		lm3		lm4		lm5	
	Estimate (Std. Error)	*t value (Pr(>|t|))*	Estimate (Std. Error)	*t value (Pr(>|t|))*	Estimate (Std. Error)	*t value (Pr(>|t|))*	Estimate (Std. Error)	*t value (Pr(>|t|))*	Estimate (Std. Error)	*t value (Pr(>|t|))*
(Intercept)	0,476 (0,054)	*8,880 (0,00)*	0,036 (0,054)	*0,672 (0,502)*	-0,040 (0,129)	*-0,313 (0,754)*	-0,090 (0,127)	*-0,712 (0,477)*	-0,060 (0,127)	*-0,470 (0,638)*
bed1	-0,426 (0,034)	*-12,557 (0,000)*	-0,342 (0,032)	*-10,586 (0,000)*	-0,341 (0,032)	*-10,556 (0,000)*	-0,303 (0,032)	*-9,497 (0,000)*	-0,295 (0,032)	*-9,270 (0,000)*
bed2	-0,341 (0,035)	*-9,822 (0,000)*	-0,406 (0,033)	*-12,265 (0,000)*	-0,402 (0,033)	*-12,162 (0,000)*	-0,372 (0,033)	*-11,387 (0,000)*	-0,363 (0,033)	*-11,134 (0,000)*
bed3	-0,214 (0,038)	*-5,600 (0,000)*	-0,528 (0,037)	*-14,270 (0,000)*	-0,522 (0,037)	*-14,104 (0,000)*	-0,501 (0,037)	*-13,722 (0,000)*	-0,494 (0,036)	*-13,552 (0,000)*
bed4	-0,205 (0,045)	*-4,584 (0,000)*	-0,651 (0,044)	*-14,946 (0,000)*	-0,643 (0,044)	*-14,755 (0,000)*	-0,621 (0,043)	*-14,446 (0,000)*	-0,617 (0,043)	*-14,388 (0,000)*
bed5 ou plus	-0,117 (0,051)	*-2,302 (0,021)*	-0,816 (0,051)	*-16,113 (0,000)*	-0,796 (0,051)	*-15,725 (0,000)*	-0,793 (0,050)	*-15,877 (0,000)*	-0,791 (0,050)	*-15,868 (0,000)*
quartier10	-0,233 (0,047)	*-4,931 (0,000)*	0,014 (0,045)	*0,319 (0,75)*	0,015 (0,045)	*0,33 (0,741)*	0,081 (0,045)	*1,817 (0,069)*	0,06 (0,045)	*1,353 (0,176)*
quartier11	-0,262 (0,046)	*-5,703 (0,000)*	0,029 (0,044)	*0,665 (0,506)*	0,029 (0,044)	*0,651 (0,515)*	0,106 (0,043)	*2,431 (0,015)*	0,078 (0,043)	*1,806 (0,071)*
quartier12	-0,241 (0,050)	*-4,794 (0,000)*	0,087 (0,048)	*1,811 (0,07)*	0,086 (0,048)	*1,797 (0,072)*	0,168 (0,048)	*3,533 (0,000)*	0,134 (0,048)	*2,824 (0,005)*
quartier13	-0,270 (0,051)	*-5,268 (0,000)*	0,068 (0,049)	*1,384 (0,166)*	0,064 (0,049)	*1,307 (0,191)*	0,144 (0,049)	*2,972 (0,003)*	0,113 (0,049)	*2,32 (0,02)*
quartier14	-0,218 (0,051)	*-4,287 (0,000)*	0,096 (0,049)	*1,967 (0,049)*	0,095 (0,049)	*1,945 (0,052)*	0,169 (0,048)	*3,506 (0,000)*	0,132 (0,048)	*2,735 (0,006)*
quartier15	-0,181 (0,047)	*-3,825 (0,000)*	0,055 (0,045)	*1,204 (0,229)*	0,054 (0,045)	*1,191 (0,234)*	0,122 (0,045)	*2,735 (0,006)*	0,097 (0,045)	*2,164 (0,03)*
quartier16	0,112 (0,051)	*2,209 (0,027)*	0,235 (0,049)	*4,838 (0,000)*	0,233 (0,048)	*4,803 (0,000)*	0,292 (0,048)	*6,098 (0,000)*	0,273 (0,048)	*5,722 (0,000)*
quartier17	-0,153 (0,048)	*-3,185 (0,001)*	0,125 (0,046)	*2,716 (0,007)*	0,122 (0,046)	*2,664 (0,008)*	0,206 (0,045)	*4,547 (0,000)*	0,181 (0,045)	*3,992 (0,000)*
quartier18	-0,235 (0,046)	*-5,149 (0,000)*	0,074 (0,044)	*1,685 (0,092)*	0,075 (0,044)	*1,711 (0,087)*	0,151 (0,043)	*3,491 (0,000)*	0,133 (0,043)	*3,066 (0,002)*
quartier19	-0,291 (0,048)	*-6,036 (0,000)*	0,093 (0,047)	*1,991 (0,046)*	0,092 (0,046)	*1,990 (0,047)*	0,190 (0,046)	*4,132 (0,000)*	0,162 (0,046)	*3,525 (0,000)*
quartier2	0,131 (0,052)	*2,521 (0,012)*	0,195 (0,049)	*3,936 (0,000)*	0,190 (0,049)	*3,840 (0,000)*	0,208 (0,049)	*4,278 (0,000)*	0,194 (0,049)	*3,981 (0,000)*
quartier20	-0,289 (0,048)	*-5,994 (0,000)*	0,105 (0,046)	*2,270 (0,023)*	0,103 (0,046)	*2,229 (0,026)*	0,196 (0,046)	*4,284 (0,000)*	0,165 (0,046)	*3,591 (0,000)*
quartier3	0,048 (0,050)	*0,957 (0,339)*	0,096 (0,047)	*2,034 (0,042)*	0,095 (0,047)	*2,004 (0,045)*	0,120 (0,047)	*2,574 (0,01)*	0,108 (0,047)	*2,315 (0,021)*
quartier4	-0,003 (0,054)	*-0,048 (0,962)*	0,017 (0,051)	*0,325 (0,745)*	0,012 (0,051)	*0,242 (0,809)*	0,018 (0,051)	*0,353 (0,724)*	0,009 (0,051)	*0,187 (0,852)*
quartier5	-0,137 (0,052)	*-2,606 (0,009)*	-0,002 (0,05)	*-0,042 (0,967)*	-0,003 (0,05)	*-0,063 (0,95)*	0,037 (0,049)	*0,747 (0,455)*	0,020 (0,049)	*0,407 (0,684)*
quartier6	0,048 (0,055)	*0,884 (0,377)*	0,059 (0,052)	*1,127 (0,26)*	0,060 (0,052)	*1,154 (0,248)*	0,069 (0,051)	*1,352 (0,176)*	0,057 (0,051)	*1,124 (0,261)*
quartier7	-0,009 (0,057)	*-0,166 (0,868)*	0,041 (0,054)	*0,752 (0,452)*	0,038 (0,054)	*0,702 (0,483)*	0,054 (0,053)	*1,007 (0,314)*	0,055 (0,053)	*1,043 (0,297)*
quartier8	0,183 (0,058)	*3,185 (0,001)*	0,196 (0,055)	*3,580 (0,000)*	0,194 (0,055)	*3,547 (0,000)*	0,232 (0,054)	*4,304 (0,000)*	0,222 (0,054)	*4,119 (0,000)*
quartier9	-0,104 (0,05)	*-2,089 (0,037)*	0,103 (0,047)	*2,163 (0,031)*	0,099 (0,047)	*2,084 (0,037)*	0,167 (0,047)	*3,564 (0,000)*	0,147 (0,047)	*3,137 (0,002)*
roomPrivate room	0,280 (0,018)	*15,795 (0,000)*	0,415 (0,018)	*23,688 (0,000)*	0,411 (0,018)	*23,425 (0,000)*	0,378 (0,017)	*21,747 (0,000)*	0,395 (0,017)	*22,69 (0,000)*
roomShared room	0,408 (0,067)	*6,091 (0,000)*	0,643 (0,064)	*10,000 (0,000)*	0,636 (0,064)	*9,901 (0,000)*	0,614 (0,063)	*9,686 (0,000)*	0,641 (0,063)	*10,134 (0,000)*
minimum_nights			-0,003 (0,005)	*-0,608 (0,543)*	-0,002 (0,005)	*-0,396 (0,692)*	0,014 (0,005)	*2,621 (0,009)*	0,018 (0,005)	*3,295 (0,001)*
reglTRUE			-0,082 (0,011)	*-7,406 (0,000)*	-0,081 (0,011)	*-7,288 (0,000)*	-0,043 (0,011)	*-3,869 (0,000)*	-0,038 (0,011)	*-3,462 (0,001)*
			0,288	*41,433*	0,280	*39,929*	0,294	*42,121*	0,301	*43,084*

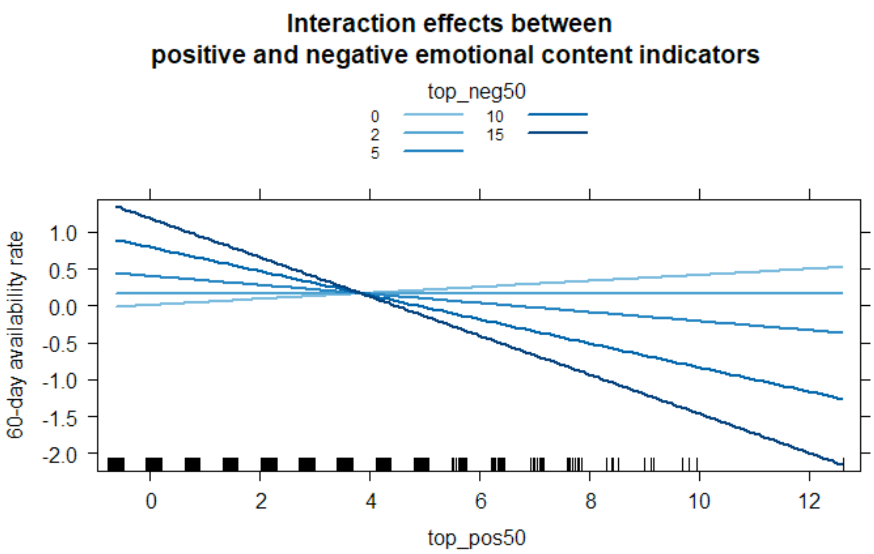
prix		(0,007)	*(0,000)*	(0,007)	*(0,000)*	(0,007)	*(0,000)*	(0,007)	*(0,000)*
		0,049	*4,002*	0,046	*3,757*	0,033	*2,755*	0,032	*2,627*
cautionTRUE		(0,012)	*(0,000)*	(0,012)	*(0,000)*	(0,012)	*(0,006)*	(0,012)	*(0,009)*
		0,151	*11,357*	0,150	*11,241*	0,145	*11,02*	0,146	*11,134*
menageTRUE		(0,013)	*(0,000)*	(0,013)	*(0,000)*	(0,013)	*(0,000)*	(0,013)	*(0,000)*
		0,099	*8,601*	0,089	*7,600*	0,036	*3,072*	0,038	*3,214*
res_instTRUE		(0,012)	*(0,000)*	(0,012)	*(0,000)*	(0,012)	*(0,002)*	(0,012)	*(0,001)*
		-0,020	*-1,401*	-0,024	*-1,664*	-0,052	*-3,738*	-0,056	*-4,034*
annulmoderate		(0,014)	*(0,161)*	(0,014)	*(0,096)*	(0,014)	*(0,000)*	(0,014)	*(0,000)*
		0,249	*16,833*	0,245	*16,544*	0,187	*12,659*	0,184	*12,458*
annulstrict		(0,015)	*(0,000)*	(0,015)	*(0,000)*	(0,015)	*(0,000)*	(0,015)	*(0,000)*
host_total_				0,025	*4,598*	0,027	*5,132*	0,028	*5,199*
Listings_count				(0,005)	*(0,000)*	(0,005)	*(0,000)*	(0,005)	*(0,000)*
				-0,088	*-8,041*	-0,104	*-9,556*	-0,104	*-9,653*
check_identityTRUE				(0,011)	*(0,000)*	(0,011)	*(0,000)*	(0,011)	*(0,000)*
				0,113	*0,956*	0,127	*1,092*	0,122	*1,050*
profile_picTRUE				(0,118)	*(0,339)*	(0,116)	*(0,275)*	(0,116)	*(0,294)*
				0,059	*4,168*	0,02	*1,37*	0,012	*0,825*
superhostTRUE				(0,014)	*(0,000)*	(0,015)	*(0,171)*	(0,015)	*(0,409)*
				0,025	*4,514*	0,013	*2,475*	0,015	*2,838*
anciennete				(0,005)	*(0,000)*	(0,005)	*(0,013)*	(0,005)	*(0,005)*
						-0,066	*-11,789*	-0,06	*-10,655*
review_scores_rating						(0,006)	*(0,000)*	(0,006)	*(0,000)*
						-0,004	*-0,786*	-0,004	*-0,83*
dernier_com						(0,005)	*(0,432)*	(0,005)	*(0,407)*
						0,150	*26,123*	0,106	*14,659*
number_of_reviews						(0,006)	*(0,000)*	(0,007)	*(0,000)*
								0,041	*5,521*
top_pos50								(0,007)	*(0,000)*
								0,079	*10,207*
top_neg50								(0,008)	*(0,000)*
								-0,020	*-7,424*
top_pos50:top_neg50								(0,003)	*(0,000)*
R^2	0,036	0,127		0,131		0,154		0,158	
Adj. R^2	0,035	0,126		0,129		0,152		0,157	
Num. Obs.	31090	31090		31090		31090		31090	
RMSE	0,982	0,935		0,933		0,921		0,918	

Interaction effect of salient positive and negative emotional content

As we can see in Table 3, the interaction between positive and negative emotional content has a significant negative effect on the availability rate of listings (β = -0.020, p-value = 0.000). Figure 2 shows the different values of the availability rate at 60 days according to the positivity indicator and according to given values of the negativity indicator (from 0 to 15). The availability rate is lowest when these two indicators are the strongest (dark blue curve). The effect of mixed emotional content on the degree of attractiveness of an offer is therefore confirmed **(H2 validated).**

Figure 2

Interaction effect between positive and negative emotional content on the 60-day availability rate



Impact of the planning level of the stay

The quality of the models increases with the time horizon studied (see Table 1). The effect of the salient sentiments contained in the reviews increases with the availability horizon, which may be related to the different short and long term booking strategies of travelers (Table 4). **H3 is therefore validated.**

Table 4

Comparison of the coefficients of the positive and negative indicators of emotional content

	lm5 15-day	lm5 30-day	lm5 60-day
top_pos50	0.033***	0.037***	0.041***
	(0.008)	(0.007)	(0.007)
top_neg50	0.073***	0.079***	0.079***
	(0.008)	(0.008)	(0.008)
top_pos50:top_neg50	-0.017***	-0.018***	-0.020***
	(0.003)	(0.003)	(0.003)
R^2	0.125	0.133	0.158
Adj. R^2	0.124	0.132	0.157
Num. obs.	31090	31090	31090
RMSE	0.936	0.932	0.918

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

Comparison of the coefficients of positive and negative emotional content indicators

Discussion

The housing offers available on peer-to-peer platforms belong, according to Nelson's typology (1970), to the category of experience goods. For these goods, whose objective characteristics alone do not necessarily allow the consumer to form an opinion, social influencing factors are expected to play a decisive role. Our research based on the reviews left on the Airbnb platform and the availability of rental housing supports this hypothesis. The introduction of features related to online reviews has helped improve the fit of the model explaining availability, as has the addition of parameters reflecting the sentiments expressed in these reviews. The results also confirm the important role of negative reviews.

A first empirical contribution of this work is to look at the behavior actually observed on the platforms (reservation or not of products) and not at intentions measured under experimental conditions. A second empirical contribution concerns the scale at which we are working in the most complete model, i.e. almost the entire corpus of reviews in French, i.e. 267,444 reviews out of the 282,057 submitted in French. A comparison of the models shows that the quality of the model improves with the number of reviews taken into account for the sentiment indicators (7, 20 or 50). A final empirical contribution is the comparison of prediction quality as a function of the prediction horizon: the longer the prediction horizon, expressed in days, the better the prediction quality, which can be explained by the more planned nature of visitor behavior.

This research shows the complex role that online reviews play in the internet user's decision-making process. The emotional content of a review has an impact on the final decision: very positive content is associated with more available offers on average, as is very negative content. Conversely, offers with reviews that express mixed emotional content (both positive and negative) are the least available. These results are all the stronger the further away the time horizon of the decision is. From a theoretical point of view, these results provide a better understanding of the impact of sentiments expressed in online reviews on consumer evaluations of an offer. Two mechanisms seem to be at work simultaneously when the internet user reads a sample of reviews. The first is the elimination of offers with converging negative reviews, as these seem too risky. These results confirm those of Kim and Gupta (2012). The second seems to go against intuition: internet users tend to reject offers with extremely positive reviews. A "too good to be true" effect could explain this phenomenon, following the example of the work of Maslowska et al. (2017). These two effects seem to show that internet users, engaged in a strategy of reading online reviews, are able to identify and distinguish the stylistic effects of the reviews and integrate them into their decision-making process.

Limitations and future research directions

This research leads to many extensions, both methodologically and conceptually. The analysis of sentiment here was based on indicators of positivity and negativity associated with the reviews left by internet users and calculated using the NRC lexicon. Future research should be interested in cross-referencing these sentiment indicators with other measures, performed by machine learning or using other sentiment lexicons (e.g., the Linguistic Inquiry and Word Count (Piolat et al., 2011) or the Lexicoder Sentiment Dictionary (Duval and Pétry, 2016)). Cross-referencing these sentiment indicators would help confirm the results obtained. Another extension would be to also integrate the topics addressed by the clients in order to see if certain topics are more related to an overall positive (or negative) sentiment. Finally, replicating the analyses on another set of data (the corpus of reviews in English, for example) would make it possible to strengthen the external validity of the results obtained.

Building on the work of Maslowska, Malthouse and Bernritter (2017), a methodological contribution of our research consists in the exploitation of a large volume of actual booking data, put into perspective with the characteristics of listings, hosts and reviews submitted online (ratings and qualitative reviews). The results support the hypothesis that reviews that are extremely positive in tone lead to lower bookings than reviews with mixed content. While the phenomenon of "too good to be true" is put forward to explain this result, only an experimental approach, measuring the psychological constructs at work in the processing of online reviews, would make it possible to confirm this mechanism. However, these approaches generally rely on a small number of reviews relating to a particular offer and cannot account for the sampling strategy used by the internet user when searching for information online. The work of Chen and Lurie (2013) thus makes it possible to highlight the mechanisms for attributing the valence of the reviews expressed and their impact on the usefulness associated with the reviews, based on a single review modified on an experimental basis (Chen and Lurie, 2013, Studies 2A, 2B, 3, 4).

The analysis carried out covered all the reviews on the Airbnb platform, without first sorting between the actual reviews and the "fakes". Several approaches have been proposed in the literature to identify fakes, based on contextual indicators (Munzel, 2015) or on the sequence of reviews for the same writer (Hu et al., 2012). In particular, based on the assumption that reviews and their writers should follow a random process, it would be possible to identify reviews from a writer who wishes to promote or destroy an offer artificially (Hu et al. 2012). It is not so much the review as the sentiments expressed by these false reviews that have a significant impact on the ranking of the promoted article. From this point of view, it would be interesting to set up a mechanism to identify false reviews and to observe, through experimentation, the reactions of consumers when they make a choice in the presence and absence of these false reviews.

The results of this research lead us to question the strategies of information processing by consumers in an information-rich environment. Future work should focus on uncovering the complex mechanisms of information selection and processing by internet users during the decision-making process.

Managerial implications

From a managerial point of view, our research shows that online reviews and the sentiments expressed therein do play a role in the user's decision process, but a weak one. Moreover, the presence of reviews with negative emotional content enhances the credibility of an offer, especially when they are associated with more positive content. These two effects tend to advise professionals offering their services or products on the platforms, as well as platform managers, to leave negative reviews visible. This will strengthen the credibility associated with the reviews and consequently the credibility associated with positive reviews, enhancing the value of the offer.

It might also be relevant to present the reviews associated with an offer in two columns side by side, one presenting negative reviews and the other positive ones, in order to help consumers make up their minds about the offer presented, while at the same time enhancing the platform's credibility by giving access to unabridged information about the quality of its offers.

Finally, we can advise that platforms set up tools for internet users to give their feedback on the usefulness of a review for their decision, which will enable them to identify the characteristics of relevant content and highlight useful reviews for consumers in order to facilitate their decision process.

References

Alkalbani A.M., Ghamry M.A., Hussain K.F. and Hussain O.K. (2016), “Sentiment analysis and classification for software as a service reviews”, *IEEE 30th International Conference on Advanced Information Networking and Applications*, 2-7.

Bansal H.S. and Voyer P.A. (2000), “Word-of-mouth processes within a services purchase decision context”, *Journal of Service Research*, 3, 2, 166-177.

Berg J., Dickhaut J. and MacCabe K. (1995), “Trust, reciprocity and social history”, *Games and Economic Behavior*, 10, 122-142.

Cabral L. and Hortacsu A. (2010), “The dynamics of seller reputation: evidence from eBay”*, Journal of Industrial Economics*, 58, 1, 54-78.

Chardon B. (2013), “Chaîne de traitement pour une approche discursive de l’analyse d’opinion”, *Doctoral Thesis in Artificial Intelligence*, Université Toulouse 3 Paul Sabatier.

Chen Z. and Lurie N.H. (2013), “Temporal contiguity and negativity bias in the impact of online word of mouth”, *Journal of Marketing Research*, 50, 4, 463-476.

Chevalier J.A. and Mayzlin D. (2006), “The effect of word of mouth on sales: online book reviews”, *Journal of Marketing Research*, 43, 345-354.

Crépin A. and Ngobo P-V. (2019), “Détermination du sentiment exprimé au sein de données textuelles issues d’Internet: utilisation d’émojis et d’un modèle neuronal”, *Proceedings of the 18th Colloque sur le Marketing Digital*, Paris.

Duval D. and Pétry F. (2016) «L’analyse automatisée du ton médiatique : construction et utilisation de la version française du Lexicoder Sentiment Dictionary», *Revue canadienne de science politique*, 49, 2, 197-220.

Fink L., Rosenfeld L. and Ravid G. (2018), “Longer online reviews are not necessarily better”, *International Journal of Information Management*, 39, 30-37.

Hong H., Xu D., Xu D., Wang G.A. and Fan W. (2017), “A meta-analysis on the determinants of online review helpfulness”, *Proceedings of the WHICEB 2017 Conference,* 34.

Hornik K, Mair P, Rauch J, Geiger W, Buchta C and Feinerer I (2013). “The textcat Package for n-Gram Based Text Categorization in R.” *Journal of Statistical Software*, 52, 6, 1-17. doi: 10.18637/jss.v052.i06.

Hu N., Bose I., Koh N.S. and Liua L. (2012), “Manipulation of online reviews: An analysis of ratings, readability, and sentiments”, *Decision Support Systems*, 52, 3, 674-684.

Kim J. and Gupta P. (2012), “Emotional expressions in online user reviews: How they influence consumers’ product evaluations”, *Journal of Business Research*, 65, 7, 985-992.

Lallement J. (2010), “L’impact de la pression temporelle sur le traitement des informations”, *Recherche et Applications en Marketing*, 25, 4, 45-68.

Lallement J. and Zollinger M. (2013), “Vite et à tout prix ?” - Ou l’importance du prix pour le consommateur pressé, *Décisions Marketing*, 69, 11-24.

Larceneux F. (2007), “Buzz et recommandations sur Internet: quels effets sur le box-office?”, *Recherche et Applications en Marketing*, 22, 3, 45-64.

Li X., Wu C. and Mai F. (2019), “The effect of online reviews on product sales: A joint sentiment-topic analysis”, *Information & Management*, 56, 2, 172-184.

Lumeau M., Masclet D. and Penard T. (2015), “Reputation and social (dis)approval in feedback mechanisms: An experimental study”, *Journal of Economic Behavior & Organization*, 112, 127-140.

Masclet D. and Pénard T. (2012), “Do reputation feedback systems really improve trust among anonymous traders? An experimental study”, *Applied Economics*, 44, 35, 4553-4573.

Maslowska E., Malthouse E.C. and Bernritter S.F. (2017), “Too good to be true: the role of online reviews’ features in probability to buy”, *International Journal of Advertising*, 36, 1, 142-163.

Mohammad, S. M. and Turney, P. D. (2013), “Crowdsourcing a word-emotion association lexicon”, *Computational Intelligence*, 29, 436-465.

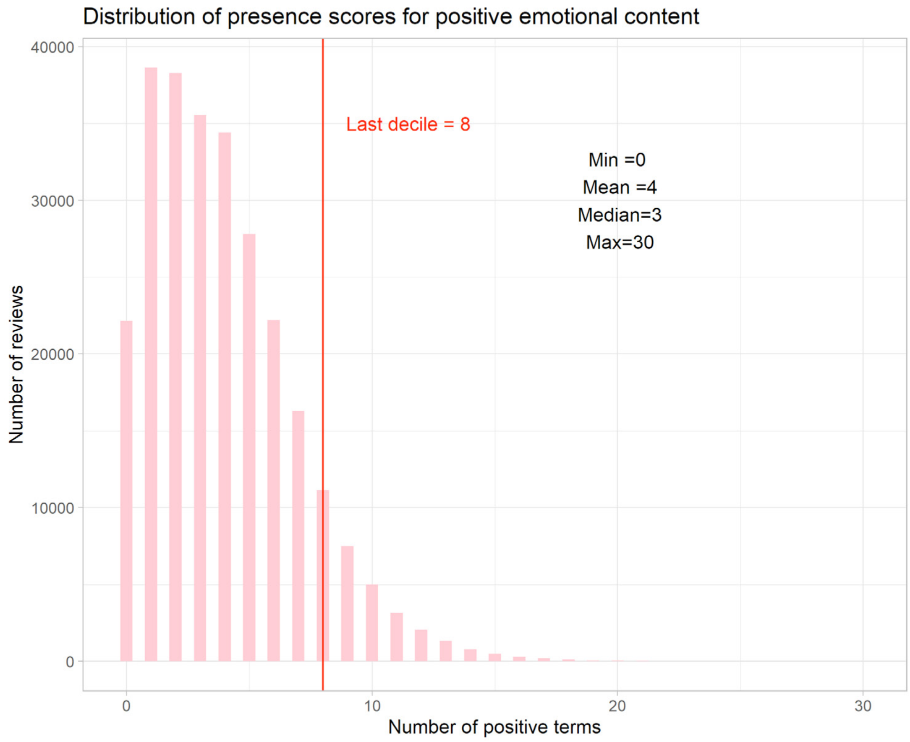
Moscarola J. and Boughzala Y. (2016), “Analyser les corpus d’avis en ligne: Analyse lexicale exploratoire et/ou modélisation sémantique ?”, *Proceedings of the 13th Journées Internationales d’Analyses Statistiques des Données Textuelles*, Nice. Disponible au lien suivant: http://lexicometrica.univ-paris3.fr/jadt/jadt2016/01-ACTES/84919/84919.pdf

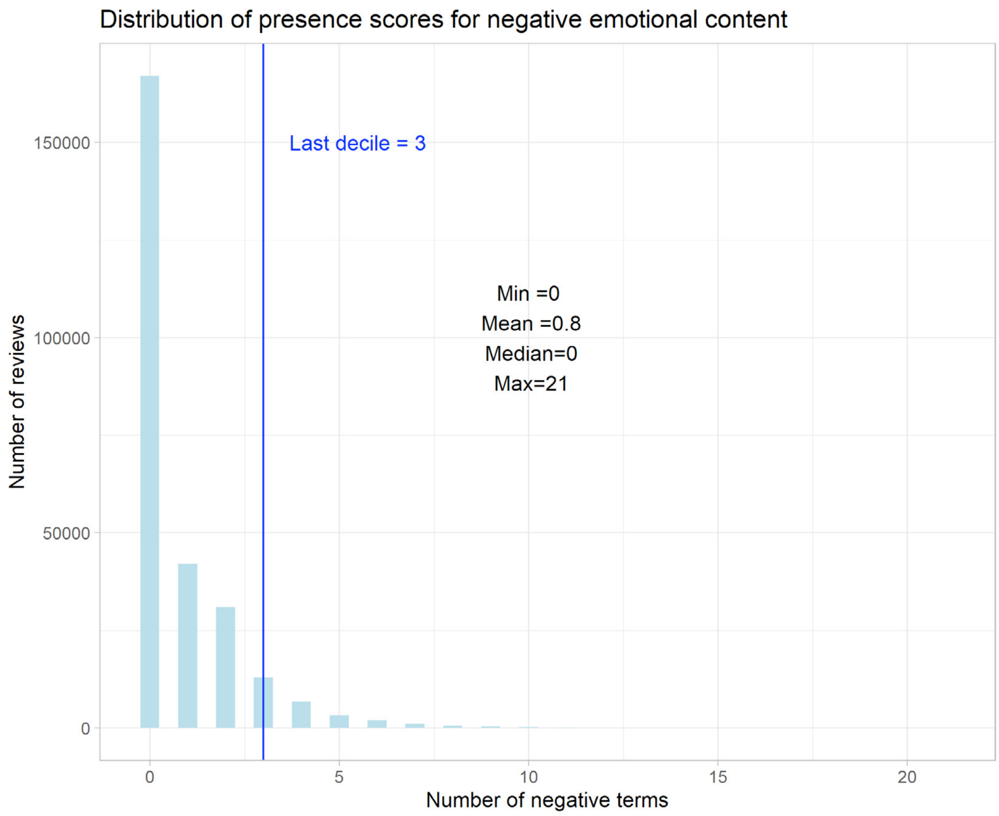
Munzel A. (2015), “Attention aux faux avis! Investigations expérimentales sur les indicateurs contextuels facilitant la détection des avis trompeurs par les internautes”, *Recherche et Applications en Marketing*, 30, 4, 24-50.

Nelson P. (1970), “Information and Consumer Behavior”, *Journal of Political Economy*, 78, 2, 311-329.
Novak P.K., Smailović J., Sluban B. and Mozetič, I. (2015), “Sentiment of Emojis”, *PLOS ONE*, 10, 12, e0144296. https://doi.org/10.1371/journal.pone.0144296.
Pavlou P.A. and Dimoka A. (2006), “The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation”, *Information Systems Research*, 17, 4, 327-414.
Pera R., Viglia G., Grazzini L. and Dalli D. (2019), “When empathy prevents negative reviewing behavior”, *Annals of Tourism Research*, 75, 265-278.
Pennebaker, J.W., Boyd, R.L., Jordan, K., and Blackburn, K. (2015), “The development and psychometric properties of LIWC2015”, Austin, TX: University of Texas at Austin.
Piolat, A., Booth, R. J., Chung, C. K., Davids, M., and Pennebaker, J. W. (2011). La version française du dictionnaire pour le LIWC: modalités de construction et exemples d’utilisation. *Psychologie française*, 56, 3, 145-159.
Ren J., Yeoh W., Shan Ee M. and Popovič A. (2018), “Online consumer reviews and sales: Examining the chicken-egg relationships”, *Journal of the Association for Information Science and Technology*, 69, 449-460.
Tran Q-V. (2015), “La crédibilité des recommandations électroniques. Etude sur l’achat de livre en ligne”, *Doctoral Thesis in Management Sciences*, Université Paris Ouest Nanterre la Défense, 30 november.
Wang W., Feng Y. and Dai W. (2018), “Topic analysis of online reviews for two competitive products using latent Dirichlet allocation”, *Electronic Commerce Research and Applications*, 29, 142-156.
Wang Y. and Chaudhry A. (2018), “When and how managers’ responses to online reviews affect subsequent reviews”, *Journal of Marketing Research*, 55, 2, 163-177.
Zhou M., Dresner M. and Windle R.J. (2008), “Online Reputation Systems: Design and Strategic Practices”, *Decision Support Systems*, 44, 4, 785-797.

Appendix

Appendix A: Distribution of presence scores of positive and negative terms in the French corpus of reviews relating to Airbnb listings in the city of Paris, as of June 6, 2019




Distribution of presence scores for negative emotional content
Last decile = 3
Min =0
Mean =0.8
Median=0
Max=21
Number of reviews
150000
100000
50000
0
0
5
10
15
20
Number of negative terms