UNDERSTANDING RELATIONSHIP QUALITY IN HOSPITALITY SERVICES: A STUDY BASED ON TEXT ANALYTICS AND PARTIAL LEAST SQUARES.

Manuel J. Sanchez-Franco, Gabriel Cepeda-Carrion and José L. Roldán
Business Administration and Marketing, University of Seville, Seville, Spain

Purpose:
The purpose of this paper is to analyze the occurrence of terms to identify the relevant topics and then to investigate the area (based on topics) of hospitality services that is highly associated with relationship quality. This research represents an opportunity to fill the gap in the current literature, and clarify the understanding of guests' affective states by evaluating all aspects of their relationship with a hotel.

Design/methodology/approach:
This research focuses on natural opinions upon which machine-learning algorithms can be executed: text summarization, sentiment analysis and latent Dirichlet allocation (LDA). Our data set contains 47,172 reviews of 33 hotels located in Las Vegas, and registered with Yelp. A component-based structural equation modeling (partial least squares (PLS)) is applied, with a dual – exploratory and predictive – purpose.

Findings:
To maintain a truly loyal relationship and to achieve competitive success, hospitality managers must take into account both tangible and intangible features when allocating their marketing efforts to satisfaction-, trust- and commitment-based cues. On the other hand, the application of the PLS predict algorithm demonstrates the predictive performance (out-of-sample prediction) of our model that supports its ability to predict new and accurate values for individual cases when further samples are added.

Originality/value:
LDA and PLS produce relevant informative summaries of corpora, and confirm and address more specifically the results of the previous literature concerning relationship quality. Our results are more reliable and accurate (providing insights not indicated in guests' ratings into how hotels can improve their services) than prior statistical results based on limited sample data and on numerical satisfaction ratings alone.

Keywords: relationship quality, Partial Least Squares, text analytics, customer reviews, exploratory and predictive analysis, Latent Dirichlet Allocation.

1. Purpose

In recent years the tourism industry has undergone an increasing number of transformation processes, caused by the development and acceptance of information technologies. Dynamic pricing, infomediation, online reservations and recommendation systems based on user-generated content (UGC) have all taken place in the tourism sector and, as a result, transformed the way in which individuals search for experiential travel services (Raguseo *et al.*, 2017; Sparks *et al.*, 2016). UGC in particular is viewed as spontaneous, insightful and passionate feedback "provided by customers that is widely available, free or low cost, and easily accessible anywhere, anytime" (Guo *et al.*, 2016, p.468) and it allows customers to describe, relive, reconstruct, and share their experiences. Consequently, UCG provides an opportunity for indirect experiences and therefore, for developing (or terminating) long-term online relationships and genuine customer loyalty. Although customer reviews are poorly structured, being more or less focused on a single entity or aspect of hospitality, or are multi-lingual, they are a major source of information for academics and managers that can help to provide a full understanding of guests' preferences and demands, and what it is that predicts whether or not they will return to a hotel.

Due to the high costs typically involved with investments in the hospitality industry, it is sensible to study the service components that customers assess, describe and share in their reviews. While star-based ratings (numerical and easily understood, with a lower search cost, and viewed as an overall assessment of the customers' post-consumption experience) are so critical that an extra half-star allows restaurants to sell out 19% more frequently (*cf.* Ghose and Ipeirotis, 2011), individuals also analyze reviews by focusing not only on the summary star ratings but also on the content of customers' *free-form text* based on subjectively-experienced intangibles (Serra-Cantallops and Salvi, 2014).

Our research analyzes a sample of 47,172 reviews of 33 urban hotels located in Las Vegas, United States of America and registered with Yelp, a review aggregator of travel-related content such as TripAdvisor and Trivago. Our research focuses on (a) pre-processing the dataset to understand the structure of the hotel review corpus; (b) identifying guest experience-related topics; (c) examining the underlying semantic structure and reducing the number of topics into meaningful groupings that makes them easier to interpret; and subsequently (d) providing an explicit representation of hotel reviews that could predict the development of long-term, mutually beneficial relationships with customers (relationship quality –RQ– *cf.* Crosby *et al.*, 1990; Dwyer *et al.*, 1987; Hennig–Thurau and Klee, 1997; Lin and Ding, 2005; Sánchez-Franco *et al.*, 2009, among others). The present study extends previous research on community services by using a marketing framework

–the RQ model– focusing on the true commitment phase (taken as the desire to maintain a relationship in the future that affectively benefits us) (Fogg and Eckles, 2007). The following questions are therefore relevant. What do guests *say* is important in the selection and evaluation of hospitality services? What are the underlying topics in hotel reviews? What form does the relationship between subjectively-experienced features and RQ assume? Can guests' experiences as represented in customer reviews be used subsequently to predict RQ?

This research is structured as follows: after this introduction, the next section (Section 2) briefly reviews the literature on the hotel guest experience and RQ. The method section (Section 3) describes the data collection and processing approach that (1) transforms free-form text into a structured form (the data cleansing process and terms extraction) that is responsive to analysis and identifies the main terms for answering the research questions; (2) identifies which features (topic communities) of a guest's stay are addressed in the review; and (3) explores the relationships between them and the extracted RQ-dimensions, to determine the predictive power of the model:

- Term selection aims to identify the most relevant terms, using term-frequency inverse document frequency scores (tf-idf approach; *cf.* Blei and Lafferty 2009; Grün and Hornik, 2011; Salton and McGill, 1986).
- Our research focuses on topic modeling (Latent Dirichlet Allocation, LDA; *cf.* Blei *et al.*, 2003) and detecting community structures (*cf.* Newman, 2006). While counting the tf-idf values is useful in knowledge extraction, examining the distribution (and association) of the topics in the documents in relation to other topics and identifying community structures (*cf.* Newman, 2006) are more powerful approaches for understanding the context of the opinions of hospitality. Topic modeling aims to select a small subset of features (topics) that minimizes the features' redundancy and maximizes their relevance; the inclusion of redundant, irrelevant and noisy features in the model building process would cause poor predictive performance and an increased computation. Another way to conceptualize extracted topics is to categorize them by communities based on the higher semantic organization of the topics, which leads to a greater understanding.
- A third goal is to explore the magnitude of the effects of the relevant community structures on RQ in the context of hospitality services, and to assess the predictive capability of the model (Henseler, 2018). In particular, our research proposes RQ as a multidimensional construct encompassing satisfaction, affective trust and affective commitment (De Wulf *et al.*, 2001; Dorsch *et al.*, 1998; Rauyruen and

Miller, 2007; Sánchez-Franco *et al.*, 2009, among others); and measures them through rating scores and sentiment analysis –based on people's evaluative judgments and affective responses to stimuli in the texts (*cf.*, Ghasemaghaei *et al.*, 2018; Heise, 1970; Pang and Lee, 2008). In order to achieve both the exploratory and predictive purposes of our research, this study uses Partial Least Squares (PLS; *cf.* Cepeda-Carrión *et al.*, 2016; Henseler, 2018; Rigdon, 2013; Ringle *et al.*, 2015; Sarstedt *et al.*, 2017; Roldán and Sánchez-Franco, 2012).

Section 4 discusses the predictive results in depth. Finally, sections 5 and 6 set out the contributions of this investigation to the literature and practice and examine its limitations. An overview of the research approach is shown in Figure 1.

*Please insert Figure 1 here *

2. Theoretical framework

Customers increasingly visit message boards, forums, or virtual communities rather than advertisements, which reflect biased realities; generate a significant amount of free online content on subjectively-experienced intangible goods or experiences; and use online review environments as key sources of information –rich in UGC. Consequently, they reduce the potential risks associated with purchase (Sparks *et al.*, 2016). Reviews therefore play a crucial role in building hotels' online reputation by leveraging the electronic word of mouth content (eWOM; *cf.* Chong *et al.*, 2018; Hennig-Thurau *et al.*, 2004; Jalilvand and Samiei, 2012) and by attracting/retaining guests in a very efficient way (*e.g.*, Gretzel and Yoo, 2008; Park *et al.*, 2007; Ye *et al.* 2011).

Our research therefore proposes a product feature-oriented approach to explore the usefulness of applying features of the guest experience (*e.g.*, location, service quality, amenities and complementary services, sleep quality and value, cleanliness aspects or staff appearance and hotel ambiance) to predict how well the whole relationship meets guests' expectations, predictions, goals, and desires. RQ is defined here as the extent to which a hospitality relationship is able to fulfill the needs of guests (*cf.* Crosby *et al.*, 1990; Hennig-Thurau and Klee 1997; Palmatier *et al.*, 2006, among others). It is also conceptualized as a multidimensional construct that includes several related facets: satisfaction, affective trust and affective commitment (*e.g.*, Sparks and Browning, 2011; Verma *et al.*, 2012).

Firstly, satisfaction is defined as the guests' perception of the extent to which their needs, goals and desires have been fully met (*cf.* Oliver 1999; also Yoon and Uysal, 2005 for a detailed review). Satisfaction –related to the service provider's performance– becomes one of the key measures of a hotel's effectiveness in outperforming others; that is, more satisfied guests have higher quality relationships with hospitality providers (*cf.* also Dorsch *et al.*, 1998). Secondly, improving the interactive features and service-related information available on online services does not necessarily guarantee customer loyalty. In this regard, although customer reviews offer star-based ratings (an overall satisfaction measure, ranging from 1, negative, to 5, positive), star ratings do not provide extended information on customer loyalty. Indeed, RQ, as a higher-order construct, is replacing customer satisfaction as a source of superior performance. Accordingly, our study proposes an evaluation of the additional effect of affective trust –related to the service provider– and focuses on integrity and benevolence; in other words, trust is based on favorable expectations regarding the intentions and behaviors of another party (Singh and Sirdeshmukh 2000; for a detailed review, Shankar *et al.* 2003). Affective trust therefore indicates that the parties in the relationship have developed an emotional bond and develop quality relationships based on the process of making promises (Dwyer *et al.*, 1987; Gronroos, 1990; Hewett and Bearden, 2001). Thirdly, although recent studies focus on the mechanisms through which UGC generates user satisfaction or trust (*e.g.*, Sánchez-Franco *et al.*, 2016; Sparks and Browning, 2011), our research also assesses the psychological sentiments through which an attitude based on the continuation of a relationship is formed (Wetzels *et al.*, 1998). Cumulative affective commitment to the relationship with the hospitality firm is thus defined as the psychological tendency to get close to others (*cf.* Shankar *et al.*, 2003).

3. Method

3.1 Data collection

Our dataset contains 47,172 reviews of 33 urban hotels, spanning the period between March 2005 and January 2017 (see Figure 2, yearly evolution, and Figure 3, seasonality-effect), contained in the *9th Yelp Dataset Challenge*. Yelp is considered to be a social networking site, belonging to the realm of social media (Ariyasriwatana *et al.*, 2014). It is the leading rating and review site for businesses in the United States of America and currently enjoys a reputation as one of the most successful websites dedicated to travel, having grown in popularity since its inception. By the end of Q3 2017, *yelpers* had written more than 142 million reviews.

The Yelp dataset initially included all types of businesses, such as restaurants, hotels, dentists, hair stylists or mechanics. The data was grouped into five main single-object types: business, review, user, check-in and tip, of which only the first three contain information about reviews and social attributes. By filtering business records whose category contains “Hotels”, reviews were identified by states and cities. As the Yelp dataset is imbalanced with regard to location, the city of Las Vegas (Nevada) was selected as a case study because of its competitive advantage in entertainment driven by tourism and gambling pleasure (Douglass and Raiento, 2004; Rowley, 2015). Las Vegas is the third largest city in the United States in terms of tourism spending and GDP contribution, and each year attracts around 40 million tourists (World Travel and Tourism Council, 2017), which is reflected in the high number of Yelp reviews. Moreover, Las Vegas brands itself as being able to generate controversial feelings that can affect tourists' perceptions (Griskevicius *et al.*, 2009).

Please insert Figure 2 here

Please insert Figure 3 here

3.2 Data cleansing process

To avoid bias, and to keep the language variable consistent across texts, our research firstly applied the textcat package based on the R 3.4.1 statistical tool to recognize English in the reviews (*cf.* also Hornik *et al.*, 2013). This returned 41,413 reviews and 33 hotels with more than 590 reviews per hotel and fewer than 1,940 –corresponding to the second and third quartiles. Our research also evaluated the readability of reviews using the reliable Flesch-Kincaid score (here 5.7), among others, to gauge the understandability of the reviews.

Our dataset transformed and converted the data into an acceptable format; from free-form text into a structured form that is suitable for analysis (*cf.* Krippendorff, 2012; Raychaudhuri *et al.*, 2002). Our research discarded punctuation, capitalization, digits and extra whitespace; it tokenized and depluralized the terms; removed selected common stop words (*e.g.*, determiners, articles, conjunctions and other parts of speech) to filter out overly common terms, which carry no useful information; and eliminated non-English characters. Furthermore, our research cleansed the corpus by omitting terms below a pre-set minimum length (< 3 characters) and reduced terms to their stem/root form using the Porter Stemming Algorithm.

3.3 Terms extraction

Our dataset comprises independent documents consisting of non-structured review texts, associated with a specific business id and user id, and in particular, contains noise and uninformative content. Term selection aims to identify the most relevant explanatory input terms to improve the performance of text analytics and to increase the comprehensibility of the results. Our research selected a dictionary whose tf-idf values are higher than the median (with the addition of the terms *bed* and *staff*, to adequately contextualize the extracted topics). A tf-idf value therefore (a) assigns a low score to terms that are either very rare or very frequent, (b) proportionally increases the number of times a term appears in the document and which is therefore (c) offset by the frequency of the term in the corpus to balance out its general popularity, providing a total vocabulary of 424 terms and 41,413 documents.

Accordingly, our research does not process texts directly, but calculates the relationships between terms and documents through a text-mining algorithm to discover hidden semantic structures in the corpus (*cf.* Blei *et al.*, 2003, among others). It builds a so-called document–term matrix (DTM) extracted from pre-processed reviews. DTM is a "*bag of words*" representation of text, and is defined as "a structured table of numbers that can in principle be analyzed using standard techniques" (Han *et al.*, 2016, p.6). The tm package was selected for its text-mining procedures (Feinerer and Hornik, 2017).

Figure 4, a summary graph based on a term-term adjacency matrix created by calculating the linear relationships between terms, shows a network structure of relatedness terms based on frequency correlations higher than 0.1. In this graphical context, our results concur with the findings of Sparks and Browning (2011). The majority of hotel reviews concern either the core functions of the hotel (*e.g.*, clean/dirty rooms or small bathrooms) or customer service (*e.g.*, interactions with staff based on how respectfully they were treated).

*Please insert Figure 4 here *

3.4. Data mining: Features and RQ extraction

Data mining is defined as a sophisticated data search capability that uses statistical algorithms to discover patterns and relationships in data (*cf.* Rygielski *et al.*, 2002). While "the tf-idf reduction has some appealing features –notably in its basic identification of sets of terms that are discriminative for documents in the collection– the approach brings a relatively small reduction in description length and reveals little in the way of inter- or intra-document statistical structure" (Blei *et al.*, 2003, p.994; Blei, 2012). The extraction and selection of distinct topics and their semantic communities are therefore essential. Likewise, star ratings and sentiment analysis (*cf.* Das and Chen, 2001; Tong, 2001; Nasukawa and Yi, 2003; Pang *et al.*, 2002; Turney, 2002; Yi *et al.*, 2003) become key tools for summarizing satisfaction, affective trust and affective commitment.

3.4.1. Extracting topics

Topic models provide an algorithmic solution to managing, organizing and annotating large unstructured text data relating to hotel reviews that identify latent patterns of term occurrence from their distribution in the corpus. Topic models are highly flexible and widely used tools that semantically identify related documents through the topics they address and are therefore essential for summarizing documents and corpora (Blei, 2012). Our research applies the probabilistic topic model of LDA (*cf.* Blei *et al.*, 2003, among others), an unsupervised (and efficient) generative probabilistic method. The topicmodel package was selected for topic modeling with the LDA (*cf.* Grün and Hornik, 2011).

By applying LDA, all reviews share the same topic set, but each review exhibits a different probabilistic mixture of those topics. LDA, like other topic modeling algorithms, is based on two outputs: a matrix of term-probabilities, which indicates for each term the probability of its belonging to each topic –the *P(term|topic)*; and a document composition matrix, which is a probability mass distribution of topic proportions within the document *P(topic|review)*. Griffiths and Steyvers (2004) suggest an α value of 50/K as the parameter of the prior distributions for the topic distribution of documents, and 0.1 as the δ value for the parameter of the prior distribution of the term across topics, using the Gibbs sampling algorithm (*cf.* Steyvers and Griffiths, 2006). K is the number of selected topics. Our research also proposes a λ relevance-value based on logarithmized parameters of the term distribution for each topic. λ relevance-values determine the specificity of the term within the topic, where λ belongs to (0, 1) (*cf.* Sievert and Shirley, 2014). $\lambda = 0$ ranks the terms according to their probability in the entire document collection. $\lambda = 1$ ranks the words according to their topic-specific probabilities. Sievert and Shirley (2014) suggest a λ value of 3/5 as an optimal value for identifying the topics associated with the top terms that are more likely to appear within that topic than in the other documents. For brevity, our research omits an additional detailed explanation, and recommends previously cited research that describes the standard process.

By analyzing the variation of statistical perplexity during topic modeling –comparable to goodness-of-fit measures for statistical models– a heuristic approach is proposed to estimate the most appropriate number of topics K, based on an n-fold cross-validation. In order to determine the ideal number of topics, our research first applied the Rate of Perplexity Change (RPC; Zhao *et al.*, 2015), and then trained and tested the algorithm using 10-fold cross-validation (on perplexity) at different values of K (from 5 to 150 topics). Zhao *et al.* (2015) conclude that the approach is stable and accurate. Lower values In particular denote more predictive power –*i.e.*, lower values of perplexity indicate a lower misrepresentation of the terms in the test documents for the trained topics. Assuming that the change point of RPC is considered to be the most appropriate number of topics, our research establishes a useful model containing 80 topics. Perplexity fluctuates when small variations indicate an acceptable fit. See Figure 5.

To correctly interpret the themes extracted, our research plots Figures 6abc (λ relevance-value = 0.60, and also 0.00 and 1.00 to assist the interpretation/comparison of topics) with the seven most relevant terms to describe the topics.

Please insert Figure 5 here

Please insert Figures 6abc here

3.4.2. Extracting topic communities

Once the term distributions for each topic had been identified, our study employed network analysis to research the bonds between topics, based on the term distribution for each one. Network analysis provides an accurate set of methods and tools to produce structures (and sub-structures), showing how topics are organized and providing a deeper understanding of a system. Nodes (topics) might be related to higher constructs yet have no causal relation with them (*cf.* Guyon *et al.*, 2017; Marsman *et al.*, 2017). Community detection was therefore employed to examine the underlying semantic structure and further reduce the number of topics into meaningful groupings, making them easier to interpret.

In particular, our research proposes the identification of communities of topics with similar connectivity patterns (Fortunato, 2010) by maximizing the modularity measure over all possible partitions (*cf.* also Brandes *et al.*, 2008). Modularity is one of the most important measures of a partition's quality (Radicchi *et al.*, 2004), being (1) the standard objective function used in network cluster analysis, and (2) the fraction of within-community edges minus the expected value of the same quantity for a randomized network (Newman and Girvan, 2004). Modularity is thus a fit-measure of the internal density of clusters compared to the external density of a network. Modularity values usually range from about 0.3 to 0.7: empirically, a value above 0.3 is a good indicator of the significant community structure in a network (Clauset *et al.*, 2004).

The igraph package was selected for community-detection based on network analysis (Grün and Hornik, 2011). By maximizing the modularity measure over all possible partitions, its value here is equal to 0.653, based on a gini correlation-index above 0.20, and p-values < 0.001. As can be seen in Figure 7, the proposed network contains 53 nodes out of the final 80 topics, and nine communities. The largest community comprises less than 20% of the vertices.

Each community is named according to the semantic space represented by the topics in the specific community. The first community, with five topics, is *hybrid*, and represents distinct groups of topics that describe different hotel guest experiences. The second community identifies the impact of the appraisal of hotel *ambiance*, based on ambient conditions in the hospitality environment (*cf.* Jani and Han, 2014). The third community refers to *food/beverage* services. The fourth community is *amenities*, relating to topics such as the microwave oven, TV, gym, spa, and other features provided by the hotel. The fifth community is associated with *staff and service* quality, "reflecting the level of staff performance, and personalization, and the interactions between staff and guests, that is, the empathy of the staff" (Sánchez-Franco *et al.*, 2016, p. 1177). The sixth community, *family friendliness*, suggests that when guests share their story about staying at a hotel with their family members, "their experience is likely to be linked with the need for a large room or attractions they want to visit" (Xiang *et al.*, 2015). Here guests employ terms associated with an overall experience based on positive descriptions. The seventh community is *night-life*, and contains topics associated with spectacles provided by the city of Las Vegas. The eighth community, containing six topics, is named hybrid, representing a wide variety of different experiences. The ninth community, with three topics, represents the *experiential* aspects of the hotel stay relating to core product (*e.g.*, dirty, clean or wet).

Please insert Figure 7 here

3.5. Extracting RQ-dimensions

Our research uses hotel ratings for the customers' assessment of/sentiments regarding the hospitality provider's overall performance in their encounters (satisfaction), and It also proposes a sentiment analysis to detect, extract and classify opinions, based on (1) approximations of a user's psychological state, demonstrating affective trust (*e.g.*, Geyskens *et al.*, 1996); and (2) affective and emotional attachments to the service, *i.e.*, affective commitment (*e.g.*, Allen and Meyer, 1990). Our proposal refers to a dictionary of opinion terms classified into categories, each expressing a specific sentiment. As Han *et al.* (2016, p.6) recommend, "these [dictionaries] are created to summarize the opinions within online customer reviews and to perform tonal analysis of social media blogs". The syuzhet package (Jockers, 2017) was chosen for the opinion-mining procedures and our research refers to two dictionaries in particular:

- The National Research Council Canada (NRC) sentiment dictionary (developed by Mohammad and Turney, 2010, 2013) estimates the presence of eight different emotions (trust –faith and integrity– among others; see Table 1) and their corresponding valences.
- AFINN: this lexical-based approach using a metric developed by Nielsen (2011) in the microblogging space, estimates that the sentiments relating to the guest's expected outcomes are a result of them enacting the behavior. AFINN contains a list of 2,477 English word forms rated for semantic orientation, ranging from -5 (negative) to 5 (positive).

Table 1 shows the descriptive statistics for the RQ dimensions. The average review star rating is 3.21 (sd = 1.41), with an average text length of 178 terms. The rating distributions are slightly skewed towards the 4- and 5-star ratings (see Table1). Although the majority of reviews give high scores –over 20,795 4- and 5-star reviews were submitted, compared to fewer than 13,269 1- and 2-star reviews– the proportion of 1- and 2-star reviews has been increasing over time (see Figure 8). The average affective trust rating (based on NRC scores) is 3.92 (sd = 3.26). The median sentiment (trust) is 3.00 with a minimum of 0.00 and a maximum of 28. There is no linear relationship between the mean of trust scores by date (adjusted R-squared = 0.03; see Figure 9). The AFINN sentiment average is 7.82 (sd= 9.80). The median sentiment (AFINN) is 7.00 with a minimum of -45 and a maximum of 80. The distribution of sentiments is skewed towards the positive end. The average sentiment by date has decreased (linearly) slightly over time (adjusted R-squared = 0.147; see Figure 9). Apparently, contrary to Raguseo's *et al.*'s (2017) conclusions, hotels have not been learning how to effectively manage online visibility through social network sites, or to improve their hospitality services, despite the increased market transparency provided by the UGC-aggregator of tourism-related content.

Please insert Table 1 here

Please insert Figure 8 here

Please insert Figure 9 here

3.6. Partial Least Squares Structural Equation Modeling: An exploratory and predictive analysis

One essential contribution of this research is its empirical development and identification of the antecedents of RQ based on UGC by applying topic-model algorithms, and the provision of an explicit representation of hotel reviews that can predict Yelp business loyalty, based on RQ.

3.6.1. Data analysis

To aid interpretation, our research uses the logarithm of topic distributions to correct a skewed distribution, and has logarithmized all topic-based variables, except for RQ scores. To analyze the relationships between constructs and their respective indicators, a composite approach was adopted. All constructs represent a mixture of aspects, combined to form new objects (Nitzl and Chin, 2017). This composite approach is appropriate since our research uses archival data, which usually lacks comprehensive substantiation in measurement theory (Sarstedt *et al.*, 2017). Latent variables were modeled as composites –formed as linear combinations of their respective indicators (Hair *et al.*, 2017a). SmartPLS 3.2.7. (Ringle *et al.*, 2015) Partial Least Squares Structural Equation Modeling (PLS-SEM) was used, a component-based structural equation modeling technique (Rigdon, 2013). The choice of PLS-SEM is appropriate as the main objectives of this study are exploration and prediction (Hair *et al.*, 2017b; Henseler, 2018; Khan *et al.*, 2018). All constructs follow a composite measurement model (Sarstedt *et al.*, 2016), and with regard to distribution data (Gefen *et al.*, 2011), our dataset does not meet the special requirements for covariance-based SEM (CBSEM) analysis.

PLS-SEM analysis is divided into three stages. The first requires the evaluation of the measurement model (the outer model), and assesses the relationships between observable variables and composite constructs. The exogenous and endogenous composites are estimated in Mode B (regression weights) (Rigdon, 2016). Given that composite 2 (hotel ambiance) consists of positive and negative topics, and assuming that, in the semantic space that represents the hospitality experience, these two topic groups belong to two different sub-contexts, our research divides composite 2 into two sub-composites (hotel ambiance 2a –positive valence; and 2b –negative valence) to improve the interpretation of results.

Our research also evaluates the structural model (inner model) to assess the sign, magnitude, and relevance of the relationships between composites (Roldán and Sánchez-Franco, 2012). The path coefficients are the most important result of the structural model. Bootstrap percentile confidence intervals of the path coefficients help to assess the relevance of the estimated parameters (Chin, 1998). Finally, in the third stage, our research assesses the predictive performance of our PLS-SEM model (Cepeda-Carrión *et al.*, 2016) using holdout samples (Shmueli *et al.*, 2016).

3.6.2 Preliminary analysis

Although our network analysis identifies nine composites (communities), a preliminary (exploratory) PLS analysis concluded that only four of these reached effect size f^2 values higher than the recommend base value of 0.02 (Cohen, 1988). These four composites were therefore retained for the final analysis since they affect the RQ variable above the minimum level. Our final exploratory and predictive model consists of four composites linked to RQ. It should be noted that the available data was used to develop, modify, and test the model, which is the essence of the exploratory feature of PLS. Consequently, the model itself could to some extent be considered as a random variable, and both the parameter estimates and their confidence intervals should be interpreted with caution since they are the result of an exploration.

3.6.3. Outer model assessment

Our research evaluates the nomological (external) validity of the composites through a confirmatory composite analysis (Henseler *et al.*, 2014). The standardized root mean square residual (SRMR) index for the saturated model was estimated (Henseler *et al.*, 2016), achieving a SRMR value of 0.0267, which is well below the usual threshold of 0.08 (Hu and Bentler, 1999). Following Henseler (2017), this means that the composites of our model would behave within a nomological net rather than as isolated manifest variables.

A critical issue for composites estimated in Mode B is the potential multicollinearity of their manifest variables. The Variance Inflation Factor (VIF) values for the range of those items are below 2.058 (see Table 2) and individual items of the composites exhibit no issues of multicollinearity. Following Chin (2010), our research reports the loadings, weights and the significance of the weights of individual items for each composite in Table 2.

Please insert Table 2 here

3.6.4. Inner model assessment

The final exploratory model explains the 38.3% variance (R^2) in RQ, and suggests an appropriate predictive power (in-sample prediction) for the dependent variable, since it is above the moderate effect value (0.33) indicated by Chin (1998). Table 3 includes the percentage variance in the dependent variable explained by each independent variable. See also Figure 10.

As Henseler *et al.* (2009) comment, the use of bootstrapping (5,000 resamples) generates *t*-statistics and confidence intervals for the standardized regression coefficients, which allows the relevance of each direct effect to be identified. In this regard, all path coefficients in Table 3 appear to be relevant for explaining the RQ variable. The four direct effects also return f^2 values above the low effect value (0.02) indicated by Chin (1998). We would highlight the key role of composite 5 (relating to staff appearance), which explains 23.19% of the variance of RQ.

Please insert Figure 10 here

Please insert Table 3 here

3.6.5. Predictive performance of the model using holdout samples

Our research assesses the predictive power of our model (out-of-sample prediction) using the PLS predict algorithm developed by Shmueli *et al.* (2016), included in SmartPLS software version 3.2.7, and conducts a cross-validation process using holdout samples. This evaluation indicates whether our model is able to generate accurate predictions of new interpretable observations (Shmueli and Koppius, 2011). Our research applies a benchmark developed by the SmartPLS team (SmartPLS, 2017), particularly the Q^2 value. This index compares the prediction errors of the PLS path model against simple mean predictions. A positive Q^2 means the prediction error of the PLS-SEM results is smaller than the prediction error simply using the mean values. In our case, the model shows a satisfactory predictive performance both for the endogenous composite (RQ) and for its manifest variables. See Table 4.

Please insert Table 4 here

Finally, our research also assesses the predictive validity of our model, focusing on the overfitting issue *i.e.*, is the model fit geared too much towards training data or will it perform comparably with new data. To answer this question, we followed the guidelines suggested by Danks *et al.* (2017), which were applied in Felipe *et al.* (2017). Thus, in-sample versus out-of-sample predictions were compared to actual composite scores.

The composite RQ resulting from this approach returned the following metrics: in-sample root mean squared error (RMSE) (IS) = 0.833151, and out-of-sample RMSE (OOS) = 0.833023. RMSE can be interpreted as a standard deviation since component scores are normalized (mean 0 and variance 1). The difference between in-sample and out-of-sample RMSE is 0.000128, which is practically zero. Given that the difference in RMSE is not substantial, overfitting is not a problem for this study. The density plots of the in-sample and out-of-sample residuals are set out in Figure 11, showing an extreme overlap.

Please insert Figure 11 here

In conclusion, both prediction analyses show that our model has sufficient predictive power (out-of-sample prediction) to predict values for a new dataset. The four composites therefore appear to predict RQ in additional samples.

4. Discussion and conclusions

Hotel firms offer essentially homogeneous services, and try to serve guests' interests by standing out from their competitors (Xiang *et al.*, 2015). Guest reviews or comments expressed in natural language allow customers to describe their (latent) opinions and experiences of hospitality services. This study is therefore not restricted to quantitative variables and consequently identifies performance issues that are subtle yet difficult to diagnose, and which may damage the hotel's reputation if left unaddressed.

The objective of our research is to extract the latent dimensions from 41,413 online guest reviews with the aims of (1) offering new insights into the determinants of a guest's affective state by evaluating all aspects of their relationship with an urban hotel; (2) understanding what makes guests return to a hotel or not, which is key to its success and long-term competitiveness; and (3) incorporating managerial results into the customer decision-making process. Urban tourism is one of the most popular forms of tourism but it has received "a disproportionately small amount of attention from scholars of either tourism or of the city" (Ashworth and Page, 2011, p.1).

The topics (and their communities) explored here and their predictive contributions to RQ are identified by applying topic-modeling algorithms, network analysis and PLS-SEM. The application of text analytics provides a summarized structure of UGC, by grouping comments into topic communities. Topic modeling is indeed an important field for summarizing and understanding ever-expanding online information archives, and contributes to tourism research. Furthermore, by using techniques for NLP and LDA as an unsupervised learning model, our analysis confirms and addresses more specifically the results of the previous literature concerning hotel features. Its results should be even more reliable and accurate than prior statistical results based solely on guests' scores (customer satisfaction) and insights obtained from traditional satisfaction surveys based on small data samples.

4.1. Theoretical implications

Our research (based on a larger, unstructured, and complex dataset) enables hotels to allocate resources according to the areas that matter to guests and in particular (1) to identify the guests' specific perceptions; (2) to identify topic communities around which guests evaluate hospitality; and (3) to explore and assess their (predictive) influence on the strength of a relationship in order to meet their guests' needs.

One of the primary proposals of this research is to seek a deeper understanding of how to construct a satisfaction–trust–commitment model in the hospitality industry by (1) exploring the semantic space that represents the hospitality experience reported by guests, (2) applying sentiment analysis techniques, and (3) applying PLS-SEM. On the one hand, the application of techniques such as LDA, sentiment extraction and network analysis are a growing area of academic research and have not yet been systematically studied. On the other hand, although textual reviews have been widely studied in the literature, there has been very little research into knowledge extraction based on this type of comment and their influence on non-economic satisfaction, customer sentiments related to affective commitment and affective trust.

4.2. Managerial implications

If a hotel company is to maintain a *truly-loyal* relationship with its guests, managers must take into account both types of features of their hotel –the tangible and intangible cues– when allocating their marketing efforts. Hotels that provide the most appropriate combination of intangible and tangible features are most likely to achieve competitive advantage.

Firstly, there is a negative connotation when guests refer to their hospitality experiences relating to composites 5, 6 and 9. The attributes or benefits based on staff experience –composite 5 (intangible cues based on the non-physical nature of services, such as interactions with friendly employees)– and professionalism are essential features in the evaluation of hotel quality, *i.e.*, the highest standardized coefficient path (in absolute terms) and consequently the most influential composite associated with RQ. Likewise, *core* tangible and experiential topics –composite 9 (based on the room's poor layout or maintenance)– become essential drivers –the second most influential standardized coefficient path. Moreover, the family-friendliness construct (composite 6, related to the need for a *special* room or attractions they want to visit, *cf.* Xiang *et al.*, 2015) has the third most influential standardized coefficient path. Events in Las Vegas, such as festivals, concerts, trade shows, conventions, and sporting events are also key marketing features in the promotion of Las Vegas hotels (and their hospitality experience), given their increasingly global ability to attract visitor spending. Tourist attractions on the Las Vegas Strip should complement their gaming offerings with activities for families that are easily accessible from the street.

eWOM makes it possible for consumers to post negative comments online, "thereby making their complaints public, and shifting the intended audience to include both the business as well as other consumers" (Zhang and Vásquez, 2014, p.63). Guests' negative opinions of *poor* hospitality service are capable of suppressing their favorable sentiments towards *good* service (Han *et al.*, 2016). Our research finds that topics with stronger negative connotations are associated with lower RQ, and those with more positive connotations have higher RQ. Conversely, negative opinions have a greater influence on RQ than positive opinions. The signs of the coefficients are quite revealing. Guests may tend to write more about staff appearance (*e.g.*, poor behavior, negative attitude, lack of knowledge or skills) and core products such as furnishings when they are dissatisfied or do not trust hotel services, and may write more about the hotel ambiance and experience when they are more satisfied or have a greater trust in the hotel services. Similarly, the negative value for the family-friendliness composite based on 'travel party' suggests that this key decision cue, represented by experience-related topics, has a negative connotation for RQ. If guests' have negative perceptions of casino gambling, sightseeing, dining, and nightlife entertainment, this creates dissatisfied guests, triggering negative hotel eWOM, and reducing the likelihood of their booking the same hotel again. To summarize, negative reviews are harmful to companies (Bambauer-Sachse and Mangold, 2011), and in fact have a greater weight in the decision-making process (Herr *et al.*, 1991). Negative reviews here are stronger, more influential, and are more difficult to resist than positive reviews.

Thirdly, composite 2a is positively related to (hotel) ambiance as a part of the sensory servicescape, associated with light, sound, smell, décor, and air quality, etc. Composite 2a could be defined as 'experience', *i.e.*, the guests' overall experience that includes positive descriptive terms, such as great, excellent, and recommend. Hotel ambiance (composite 2a) influences interactions with the exchange partner –which create fulfilling, gratifying, and easy interactions– and increases guests' desire to stay (*cf.* Simpeh *et al.*, 2011). This in turn generates positive reviews about the service (Jani and Han, 2014).

Other features of the hotel that are reported as essential include the following (ordered from the highest to lowest influence on RQ):

- Composite 5: Relational (operational) aspects based on staff and service quality, reflecting the level of staff performance; personalization; and demonstrating the interactions between staff and guests, *i.e.*, staff empathy (friendliness, respectful behavior and treatment, or understanding the customer).

- Composite 9: Core (strategic and operational) services, *i.e.*, tangible factors (*e.g.*, bed, carpet) based on interior furnishings and experiential aspects such as cleanliness/dirtiness.

- Composite 6: Emphasis on family entertainment and multi-faceted tourism *i.e.*, gaming and non-gaming groups such as convention delegates and family holidaymakers.

- Composite 2a: Comfortable hotels, including size and décor. This relates to tangible factors based on hotel ambiance –how well-equipped the rooms are and the design within hotels– or intangible (experiential) cues, based on ambient conditions in the hotel environment.

While some cues such as staff, core product, festivals, concerts or trade shows, are essential for the guests' experience, others do not have a significant impact on the semantic space related to their experience. Although location is highly relevant when choosing a hotel (*cf.* Radojevic *et al.*, 2015), our research reveals only a minimal (explicit) influence of hotel location on RQ (*cf.* also Sánchez-Franco *et al.*, 2016). One possible explanation is that guests may have already considered the hotel's location when booking a room and it is not therefore explicitly elicited when evaluating their hospitality experience. Moreover, price is no longer an essential antecedent here when guests select a hotel.

In summary, as Smith *et al.* (1999) note, service failures are perceived as losses and receive a more negative weighting from customers. As Clemons and Gao (2008) suggest, this is because guests are seeking reasonable rather than optimal levels of satisfaction and the negative comments suggest that the quality threshold will not be met. This is a perceptual bias in which negative content concerning attractions at the destination carries more weight and has a greater effect on an impression than other positive evaluations. Negative comments related to room style and negative staff treatment also have an important (highest) (predictive) impact (based on f^2; *cf.* Cohen, 1988). As Xiang *et al.* (2015, p.122) note, "hygiene factors like cleanliness and maintenance do not positively contribute to satisfaction, although dissatisfaction results from their absence". Han *et al.* (2016, p.16) note that it could be better for hotels "to provide guests with a moderately good overall experience (…), because in terms of ratings the weight of the terrible service will swamp the good feelings from the stay's excellent aspects".

5. Limitations

Firstly, our research does not analyze the infrequent terms in the long tail of the distribution (*cf.*, tf-idf). A second limitation lies in the self-selection bias when guests post their reviews. For instance, "disappointed guests may have a stronger impulse to publicly share their impressions than satisfied customers" (Sánchez-Franco *et al.*, 2016, p.1183). As our research notes in Section 3.3.3, the proportion of reviews with 1- and 2-star reviews has been increasing over time. Thirdly, a topic-network structure may also change over time. In particular, latent topics represent a dynamic that potentially shifts across themes, and reviews are not written independently of each other (*e.g.*, Piramuthu *et al.* 2012; Dellarocas 2006). Fourthly, our research is focused on reviews of urban hotels in Las Vegas and may reflect certain biases of guests who visit Las Vegas. For instance, the relative importance of hotel features may be influenced by the location-based environment. Moreover, average star scores may be affected by a culture-conditioned response style (Dolnicar and Grün, 2007). Fifthly, our research relies only on reviews from Yelp and the analyzed dataset may contain some invisible bias (Han *et al.*, 2016).

6. References

Allen, N.J., and Meyer, J.P. (1990), "The measurement and antecedents of affective, continuance and normative commitment to the organization", *Journal of Occupational Psychology* Vol. 63, No. 1, pp. 1–18.

Ariyasriwatana, W., Buente, W., Oshiro, M. and Streveler, D. (2014), "Categorizing health-related cues to action: using Yelp reviews of restaurants in Hawaii", *New Review of Hypermedia and Multimedia*, Vol. 20 No. 4, pp. 317-340.

Ashworth, G. and Page, S.J. (2011), "Urban tourism research: Recent progress and current paradoxes", *Tourism Management*, Vol. 32, pp. 1–15.

Bambauer-Sachse, S. and Mangold, S. (2011), "Brand equity dilution through negative online word-of-mouth communication", *Journal of Retailing and Consumer Services*, Vol. 18 No 1, pp. 38–45.

Blei, D.M. and Lafferty, J.D. (2009), "Topic models", in Srivastava, A.N. and M. Sahami (Eds.), *Text Mining: Classification, Clustering, and Applications*, Taylor and Francis, London, pp. 71–94.

Blei, D.M., (2012), "*Probabilistic topic models*", *Communications of the ACM*, Vol. 55 No.4, pp.77.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022.

Brandes U., Delling D., Gaertler M., Görke R., Hoefer M., Nikoloski Z. and Wagner D. (2008), "On modularity clustering", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 20 No. 2, pp. 172–188.

Cepeda-Carrión, G., Henseler, J., Ringle, C.M. and Roldán, J.L. (2016), "Prediction-oriented modeling in business research by means of PLS path modeling: Introduction to a JBR special section", *Journal of Business Research*, Vol. 69 No. 10, pp. 4545–4551.

Chin, W.W. (1998), "The Partial Least Squares Approach to Structural Modeling", in Marcoulides, G.A. (ed.), *Modern Methods for Business Research*, Lawrence Erlbaum, Mahwah, NJ, pp. 295–336.

Chin, W.W. (2010), "How to Write Up and Report PLS Analyses", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds.), *Handbook of Partial Least Squares*, Springer-Verlag, Heidelberg, pp. 655–690.

Chong, A.Y.L., Khong, K.W., Ma, T., McCabe, S. and Wang, Y. (2018), "Analyzing key influences of tourists' acceptance of online reviews in travel decisions", *Internet Research*, Vol. 28, No 3, pp. 564-586.

Clauset, A., Newman, M. E. J. and Moore, Cr. (2004), "Finding community structure in very large networks", *Physical Review* E 70, 066111.1-066111.6

Clemons, E. and Gao, G. (2008), "Consumer informedness and diverse consumer purchasing behaviors: Traditional mass-market, trading down, and trading out into the long tail", *Electronic Commerce Research and Applications*, Vol. 7, pp. 3–17.

Cohen, J. (1988), *Statistical power analysis for behavioral sciences*, Erlbaum, Hillsdale, NJ.

Crosby, L. A., Evans, K.R. and Cowles, D. (1990), "Relationship quality in services selling: An Interpersonal Influence perspective", *Journal of Marketing*, Vol. 54, pp. 68-81.

Danks, N., Soumya, R. and Shmueli, G. (2017), Evaluating the Predictive Performance of Composites in PLS Path Modeling, available at: SSRN: https://ssrn.com/abstract=3055222 (accessed 25 December 2017).

Das, S. and Chen, M. (2001), "Yahoo! for Amazon: Extracting market sentiment from stock message boards," in *Proceedings of the Asia Pacific Finance Association Annual Conference* (APFA), 2001.

De Wulf, K., Odekerken-Schröder, G. and Iacobucci, D. (2001), "Investments in consumer relationships: A cross-country and cross-industry exploration", *Journal of Marketing*, Vol. 65, pp.33– 50.

Dellarocas, C. 2006, "Strategic manipulation of Internet opinion forums: implications for consumers and firms", *Management Science*, Vol. 52, No 10, pp.1577-1593.

Dolnicar, S. and Grün, B. (2007), "Assessing analytical robustness in cross-cultural comparisons", *International Journal of Tourism, Culture, and Hospitality Research*, Vol 1 No 2, pp.140–160.

Dorsch, M.J., Swanson, S.R. and Kelley, S.W. (1998), "The role of relationship quality in the stratification of vendors as perceived by customers," *Journal of the Academy of Marketing Science*, Vol. 26 No. 2, pp. 128-42.

Douglass, W. and Raiento, P. (2004), "The tradition of invention: conceiving Las Vegas", *Annals of Tourism Research*, Vol. 31 No. 1, pp. 7–23.

Dwyer, F.R., Schurr, P.H. and Oh, S. (1987), "Developing buyer-seller relationships", *Journal of Marketing*, Vol. 51, No 2, pp. 11-27.

Feinerer, I. and Hornig, K. (2017), "*A framework for text mining applications within R*", available at: https://cran.r-project.org/web/packages/tm/tm.pdf (accessed 20 October 2017).

Felipe, C.M., Roldán, J.L. and Leal-Rodríguez, A.L., (2017), "Impact of Organizational Culture Values on Organizational Agility", *Sustainability*, Vol. 9, 2354.

Fogg, B.J. and Eckles, D. (2007), "The behavior chain for online participation: how successful web services structure persuasion", in DeKort, Y., Ijsselstein, W., Midden, C.J.H., Eggen, B. and Fogg, B.J. (Eds.), *PERSUASIVE*, 2007, Springer, LNCS 4744, pp. 199-209.

Fortunato, S. (2010), "Community detection in graphs", *Physics Reports*, Vol. 486 No. 3–5, pp. 75-174

Gefen, D., Rigdon, E.E. and Straub, D. (2011), "An Update and Extension to SEM Guidelines for Administrative and Social Science Research.", *MIS Quarterly*, Vol. 35 No. 2, p. iii-A7.

Geyskens, I., Steenkamp, J.-B.E.M., Scheer, L.K. and N. Kumar (1996), "The effects of trust and interdependence on relationship commitment: A transatlantic study", *International Journal of Research in Marketing*, Vol. 13 No. 4, 303–17.

Ghasemaghaei, M., Eslami, S.P., Deal, K. and Hassanein, K. (2018), "Reviews' length and sentiment as correlates of online reviews' ratings", *Internet Research*, Vol. 28 No 3, pp.544-563.

Ghose, A. and Ipeirotis, P. (2011), "Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23 No. 10, pp. 1498–1512.

Gretzel, U., and Yoo, K. H. (2008), "Use and impact of online travel reviews". In Information and communication technologies in tourism 2008, ed. P. O'Connor, W. Höpken, and U. Gretzel, 35-46. Vienna, Austria: Springer.

Griffiths, T.L. and Steyvers, M. (2004), "Finding scientific topics", in *Proceedings of the National Academy of Sciences*, 2004, 101 (suppl. 1), pp. 5228–5235.

Griskevicius, V., Goldstein, N.J., Mortensen, C.R., Sundie, J.M., Cialdini, R.B. and Kenrick, D.T. (2009), "Fear and loving in Las Vegas: Evolution, emotion, and persuasion", *Journal of Marketing Research*, Vol. 46 No. 3, pp. 384-395.

Gronroos, C. (1990), *Service management and marketing: managing moments of truth in service competition*, Lexington Books, UK.

Grün, B. and Hornik, K. (2011), "Topicmodels: An R package for fitting topic models", *Journal of Statistical Software*, Vol. 40 No.13, pp.1-30.

Guo, Y., Barnes, S.J. and Jia, Q. (2016), "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using Latent Dirichlet Allocation", *Tourism Management*, Vol. 59, pp. 467-483.

Guyon H., Falissard B. and Kop J.-L. (2017), "Modeling psychological attributes in psychology – An epistemological discussion: Network analysis vs. latent variables", *Frontiers in Psychology*, Vol. 8, p. 798.

Hair, J.F.J., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2017a), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 2nd ed., SAGE Publications, Thousand Oaks, CA, available at: https://uk.sagepub.com/en-gb/eur/a-primer-on-partial-least-squares-structural-equation-modeling-pls-sem/book244583.

Hair, J.F., Matthews, L.M., Matthews, R.L. and Sarstedt, M. (2017b), "PLS-SEM or CB-SEM: updated guidelines on which method to use", *International Journal of Multivariate Data Analysis*, Vol. 1 No. 2, pp. 107–123.

Han, H.J., Mankad, S., Gavirneni, N. and Verma, R. (2016), "What guests really think of your hotel: Text analytics of online customer reviews", *Cornell Hospitality Report*, Vol. 16 No 2, pp. 3-17.

Heise, D.R. (1970), "The semantic differential and attitude research," in Attitude Measurement, edited by Gene F. Summers, 235–53. Chicago: Rand McNally.

Hennig-Thurau, T. and Klee, A. (1997), "The impact of customer satisfaction and relationship quality on customer retention: A critical reassessment and model development," *Psychology and Marketing*, Vol. 14 No. 8, pp. 737-64.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G. and Gremler, D. D. (2004), "Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?" *Journal of Interactive Marketing*, Vol. 18, pp. 38–52.

Henseler, J. (2017), "Bridging Design and Behavioral Research With Variance-Based Structural Equation Modeling", *Journal of Advertising*, Vol. 46 No. 1, pp. 178–192.

Henseler, J. (2018), "Partial least squares path modeling: Quo vadis?", *Quality & Quantity*, Vol. 52 No. 1, pp. 1–8.

Henseler, J., Dijkstra, T.K., Sarstedt, M., Ringle, C.M., Diamantopoulos, Straub, D.W., Ketchen, D.J., *et al.* (2014), "Common Beliefs and Reality About PLS: Comments on Ronkko and Evermann (2013)", *Organizational Research Methods*, Vol. 17 No. 2, pp. 182–209.

Henseler, J., Hubona, G. and Ray, P.A. (2016), "Using PLS path modeling in new technology research: updated guidelines", *Industrial Management & Data Systems*, Vol. 116 No. 1, pp.2-20.

Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009), "The use of partial least squares path modeling in international marketing", in Henseler, J., Ringle, C.M. and Sinkovics, R.R. (Eds.), *Advances in International Marketing*, Vol. 20, Emerald Group Publishing, Bingley, pp. 277–319.

Herr, P.M., Kardes, F.R. and Kim, J. (1991), “Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective”, *Journal of Consumer Research*, Vol. 17 No 4, pp. 454–462.

Hewett, K. and Bearden, W. O. (2001), “Dependence, trust, and relational behavior on the part of foreign subsidiary marketing operations: Implications for managing global marketing operations”, *Journal of Marketing*, Vol. 65 No 4, pp.51-66.

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*, Vol. 52 No. 6, pp.1-17.

Hu, L. and Bentler, P.M. (1999), “Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives”, *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 6 No. 1, pp. 1–55.

Jani, D. and Han, H. (2014) 'Personality, satisfaction, image, ambience, and loyalty: Testing their relationships in the hotel industry', *International Journal of Hospitality Management*, Vol. 37, pp.11-20.

Jockers, M. (2017), “*Extracts Sentiment and Sentiment-Derived Plot Arcs from Text*”, available at: https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf (accessed 20 December 2017).

Khan, G.F., Sarstedt, M., Shiau, W.L., Hair, J.F., Ringle, C.M. and Fritze, M. (2018), “Methodological research on partial least squares structural equation modeling (PLS-SEM): An analysis based on social network approaches”, *Internet Research*, forthcoming.

Krippendorff, K. (2012), *Content analysis: An introduction to its methodology*, Sage, CA.

Lin, C.P. and Ding, C.G. (2005), “Opening the black box: assessing the mediating mechanism of relationship quality and the moderating effects of prior experience in ISP service”, *International Journal of Service Industry Management*, Vol. 16 No. 1, pp. 55-80.

Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., van Bork, R., Waldorp, L.J., van der Maas, H. L. J. and Maris, G. (2017), “An introduction to network psychometrics: Relating Ising network models to item response theory models”, *Multivariate Behavioral Research*, Vol. 0 No. 0, pp. 1-21.

Mohammad, S. and Turney, P. (2010), "Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon”, in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 2010, June, LA, California.

Mohammad, S. M. and Turney, P. D. (2013), “Crowdsourcing a word–emotion association lexicon”, *Computational Intelligence*, Vol. 29 No 3, 436-465.

Nasukawa, T. and Yi, J. (2003), “Sentiment analysis: Capturing favorability using natural language processing”, in *Proceedings of the Conference on Knowledge Capture* (K-CAP), 2003.

Newman M.E.J. and Girvan M. (2004), “Finding and evaluating community structure in networks”, *Physical Review E*, Vol. 69 No. 2, pp. 026,113

Newman, M.E.J. (2006), “Modularity and community structure in networks”, PNAS, Vol. 103 No 23, pp. 8577– 8582.

Nielsen, F. A. (2011), “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs”, in *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, 2011, Vol. 718, pp. 93–98.

Nitzl, C. and Chin, W.W. (2017), “The case of partial least squares (PLS) path modeling in managerial accounting research”, *Journal of Management Control*, Springer Berlin Heidelberg, available at:https://doi.org/10.1007/s00187-017-0249-6.

Oliver R. (1999), “Whence consumer loyalty?” *Journal of Marketing*, Vol. 63 No 4, pp.33–44.

Palmatier, R. W., Dant, R. P., Grewel, D. and Evans, K. (2006), “Factors influencing the effectiveness of relationship marketing: A meta analysis”, *Journal of Marketing*, Vol. 70, pp. 136–153.

Pang, B., and Lee, L. (2008). “Opinion mining and sentiment analysis”, *Foundations and trends in information retrieval* Vol. 2 No. 1-2, pp. 1-135.

Pang, B., Lee, L. and Vaithyanathan, S. (2002), “Thumbs up? Sentiment classification using machine learning techniques”, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP, 2002, pp. 79–86.

Park, D.H., Lee, J. and Han, I. (2007), “The effect of online consumer reviews on consumer purchasing intention: The moderating role of involvement”, *International Journal of Electronic Commerce* Vol.11 No.4, pp. 125-48.

Piramuthu, S., Kapoor, G., Zhou, W. and Mauw, S. (2012), “Input online review data and related bias in recommender systems”, *Decision Support Systems*, Vol. 53, No 3, pp.418-424.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V. and Parisi, D. (2004), “Defining and identifying communities in networks”, in *Proceedings of the National Academy of Sciences*, 2004, Vol. 101, pp. 2658–2663.

Radojevic, T., Stanisic, N. and Stanic, N. (2015), “Ensuring positive feedback: Factors that influence customer satisfaction in the contemporary hospitality industry”, *Tourism Management*, Vol. 51, pp. 13–21.

Raguseo, E., Neirotti, P. and Paolucci, E. (2017), “How small hotels can drive value their way in infomediation. The case of ‘Italian hotels vs. OTAs and TripAdvisor”, *Information & Management*, Vol. 54 No. 6, pp. 745-756.

Rauyruen, P. and K. Miller (2007), "Relationship Quality as a Predictor of B2B Customer Loyalty," Journal of Business Research, Vol. 60, pp.21-31.

Raychaudhuri, S., Schütze, H. and Altman, R. B. (2002), “Using text analysis to identify functionally coherent gene groups”, *Genome Research*, Vol. 12 No. 10, pp. 1582-1590.

Jalilvand, M.R. and Samiei, N. (2012), "The impact of electronic word of mouth on a tourism destination choice: Testing the theory of planned behavior (TPB)", *Internet Research*, Vol. 22 No 5, pp. 591-612.

Rigdon, E.E. (2013), “Partial least squares path modeling”, in Hancock, G.R. and Mueller, R.O. (Eds.), *Structural Equation Modeling: A Second Course*, 2nd ed., Information Age, Charlotte, NC, pp. 81–116.

Rigdon, E.E. (2016), “Choosing PLS path modeling as analytical method in European management research: A realist perspective”, *European Management Journal*, Elsevier Ltd, Vol. 34 No. 6, pp. 598–605.

Ringle, C.M., Wende, S. and Becker, J. (2015), “SmartPLS 3”, SmartPLS, Bönningstedt, available at: http://www.smartpls.com.

Roldán, J.L. and Sánchez-Franco, M.J. (2012), “Variance-Based Structural Equation Modeling: Guidelines for Using Partial Least Squares in Information Systems Research”, in Mora, M., Gelman, O., Steenkamp, A. and Raisinghani, M. (Eds.), *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*, IGI Global, Hershey PA, pp. 193–221.

Rowley, R.J. (2015), “Multidimensional community and the Las Vegas experience”, *GeoJournal*, Vol. 80 No. 3, pp. 393-410.

Rygielski, C., Wang, J.-C. and Yen, D.C. (2002), “Data mining techniques for customer relationship management”, *Technology in Society*, Vol. 24, pp. 483–502.

Salton, G. and McGill M.H. (1986), *Introduction to Modern Information Retrieval*, McGraw-Hill, New York.

Sánchez-Franco, M.J., Villarejo-Ramos, A.F. and Martín-Velicia, F.A. (2009), “The moderating effect of gender on relationship quality and loyalty toward Internet Service Providers”, *Information & Management*. Vol. 46 No. 3, pp. 196-202.

Sánchez-Franco, M.J., Navarro-García, A. and Rondán-Cataluña, J. (2016), "Online customer service reviews in urban hotels: A data mining approach", *Psychology and Marketing*, Vol. 33 No. 2, pp. 1174-1186.

Sarstedt, M., Hair, J.F., Ringle, C.M., Thiele, K.O. and Gudergan, S.P. (2016), “Estimation issues with PLS and CBSEM: Where the bias lies!”, *Journal of Business Research*, Vol. 69 No. 10, pp. 3998–4010.

Sarstedt, M., Ringle, C.M. and Hair, J.F. (2017), “Partial Least Squares Structural Equation Modeling”, *Handbook of Market Research*, Vol. 26, Springer International Publishing, Cham, pp. 1–40.

Serra-Cantallops, A. and Salvi, F. (2014), “New consumer behavior: A review of research on eWOM and hotels”, *International Journal of Hospitality Management*, Vol. 36, pp.41-51.

Shankar, V., Smith, A.K. and Rangaswamy, A. (2003), “Customer satisfaction and loyalty in online and offline environments”, *International Journal of Marketing Research*, Vol. 20 No 2, pp.153–75.

Shmueli, G. and Koppius, O. (2011), “Predictive Analytics in Information Systems Research”, *MIS Quarterly*, Vol. 35 No. 3, pp. 553–572.

Shmueli, G., Ray, S., Velasquez Estrada, J.M. and Chatla, S.B. (2016), “The elephant in the room: Predictive performance of PLS models”, *Journal of Business Research*, Elsevier Inc., Vol. 69 No. 10, pp. 4552–4564.

Sievert, C. and Shirley, K.E. (2014), “LDAvis: A method for visualizing and interpreting topics”, in *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, 2014, Baltimore, Maryland, USA, June, pp. 63–70.

Simpeh, K.N., Simpeh, M., Abdul-Nasiru, I. (2011), “Servicescape and customer patronage of three star hotels in Ghana's metropolitan city of Accra”, *European Journal of Business and Management*, Vol. 3 No 4, pp. 119-131.

Singh, J. and Sirdeshmukh, D. (2000), “Agency and trust mechanisms in consumer satisfaction and loyalty judgments”, *Journal of Academy of Marketing Science*, Vol. 28 No 1, pp.150–67.

SmartPLS. (2017), “PLS Predict | SmartPLS”, available at: https://www.smartpls.com/documentation/ algorithms-and-techniques/predict (accessed 10 October 2017).

Smith, A. K., Bolton, R. N. and Wagner, J. (1999), “A model of customer satisfaction with service encounters involving failure and recovery”, *Journal of Marketing Research*, Vol. 36, pp. 356–372.

Sparks, B. A. and Browning, V. (2011), “The impact of online reviews on hotel booking intentions and perception of trust”, *Tourism Management*, Vol. 32, pp. 1310-1323.

Sparks, B.A., Kam Fung So, K. and Bradley, G.L. (2016), “Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern”, *Tourism Management*, Vol. 53, pp. 74-85

Steyvers, M. and Griffiths, T. (2006), “Probabilistic topic models”, in T. Landauer, D. Mcnamara, S. Dennis and W. Kintsch (Eds.), *Latent Semantic Analysis: A road to meaning*, Laurence Erlbaum Associates, Mahwah, New Jersey, pp.427-448.

Tong, R.M. (2001), “An operational system for detecting and tracking opinions in on-line discussion”, in *Proceedings of the Workshop on Operational Text Classification*, OTC, 2001.

Turney, P. (2002), “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” in *Proceedings of the Association for Computational Linguistics*, ACL, 2002, pp. 417–424.

Verma, R., Stock, D. and McCarthy, L. (2012), “Customer preferences for online, social media, and mobile innovations in the hospitality industry”, *Cornell Hospitality Quarterly*, Vol. 53 No 3, pp. 183–186.

Wetzels, M., de Ruyter, K. and van Birgelen, M. (1998), "Marketing service relationships: The role of commitment," *Journal of Business & Industrial Marketing*, Vol. 13 No 4/5, pp.406-23.

Xiang, Z., Schwartz, Z., Gerdes, J. and Uysal, M. (2015), “What can big data and text analytics tell us about hotel guest experience and satisfaction?” *International Journal of Hospitality Management*, Vol. 44, pp. 120–130.

Ye, Q., Law, R., Gu, B. and Chen, W. (2011), “The influence of user-generated content on traveler behavior: An empirical investigation on the effects of eWord-of-Mouth to hotel online bookings”, *Computers in Human Behavior*, Vol. 27 No. 2, pp. 634-39.

Yi, J., Nasukawa, T., Bunescu, R. and Niblack, W. (2003), “Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques,” in *Proceedings of the IEEE International Conference on Data Mining*, ICDM, 2003.

Yoon Y. and Uysal M. (2005), “An examination of the effects of motivation and satisfaction on destination loyalty: a structural model”, *Tourism Management*, Vol. 26 No 1, pp.45–56.

Zhang, Y. and Vásquez, C. (2014), “Hotels׳ responses to online reviews: Managing consumer dissatisfaction”, *Discourse, Context & Media*, Vol. 6, pp. 54-64.

Zhao, W., Chen, J.J., Perkins, R., Liu, Z., Ge, W., Ding, Y. and Zou, W. (2015), “A heuristic approach to determine an appropriate number of topics in topic modelling”, *BMC Bioinformatics*, Vol. 16 (suppl. 13), S8, pp.1-10.

Figure 1. Strategy for transforming free-form text into a structured form.

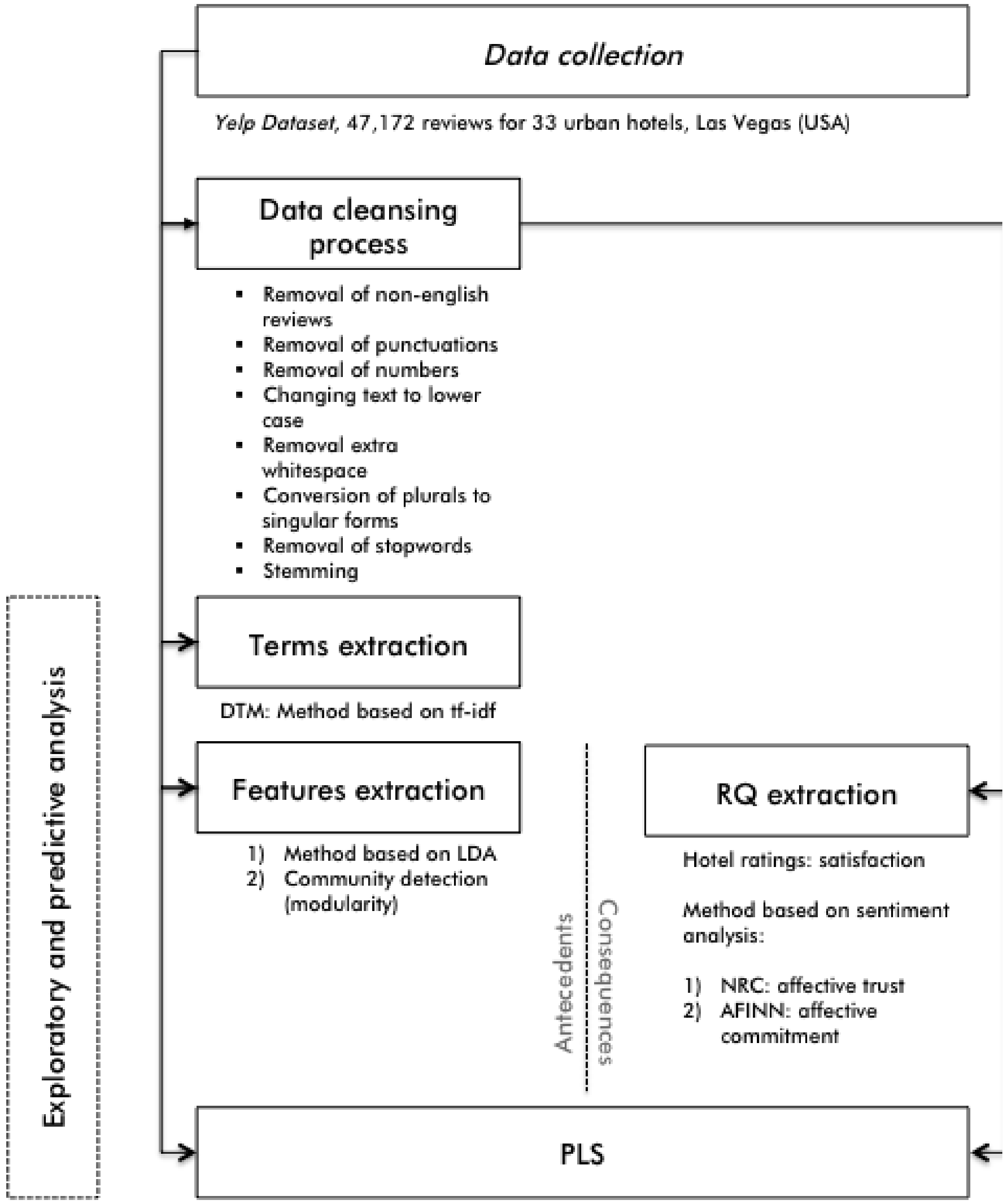


Figure 2. Evolution of reviews by date (March 2005-January 2017).

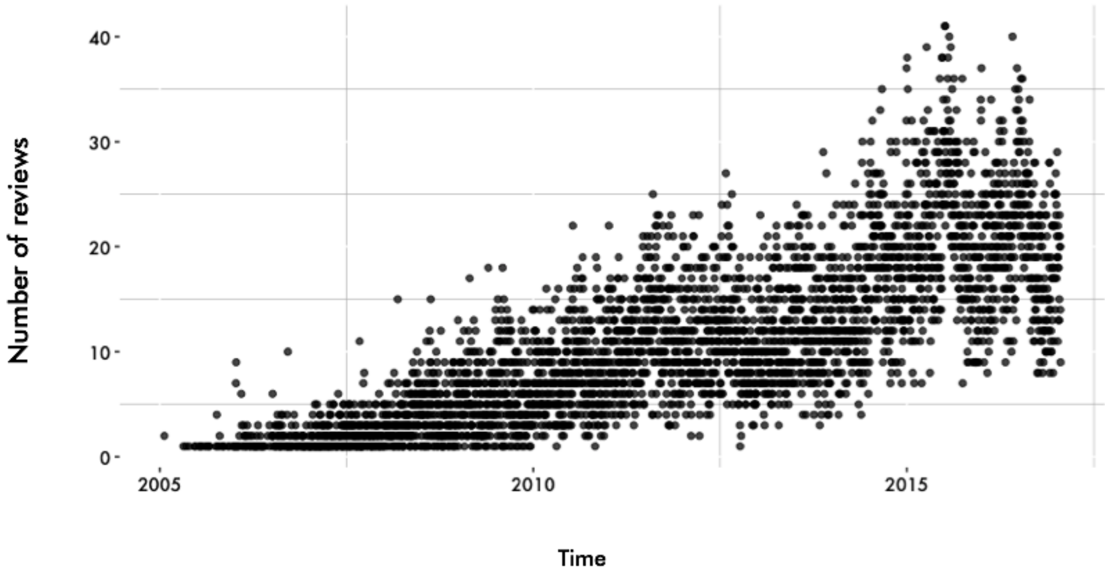


Figure 3. Seasonality effect by month-quarters.

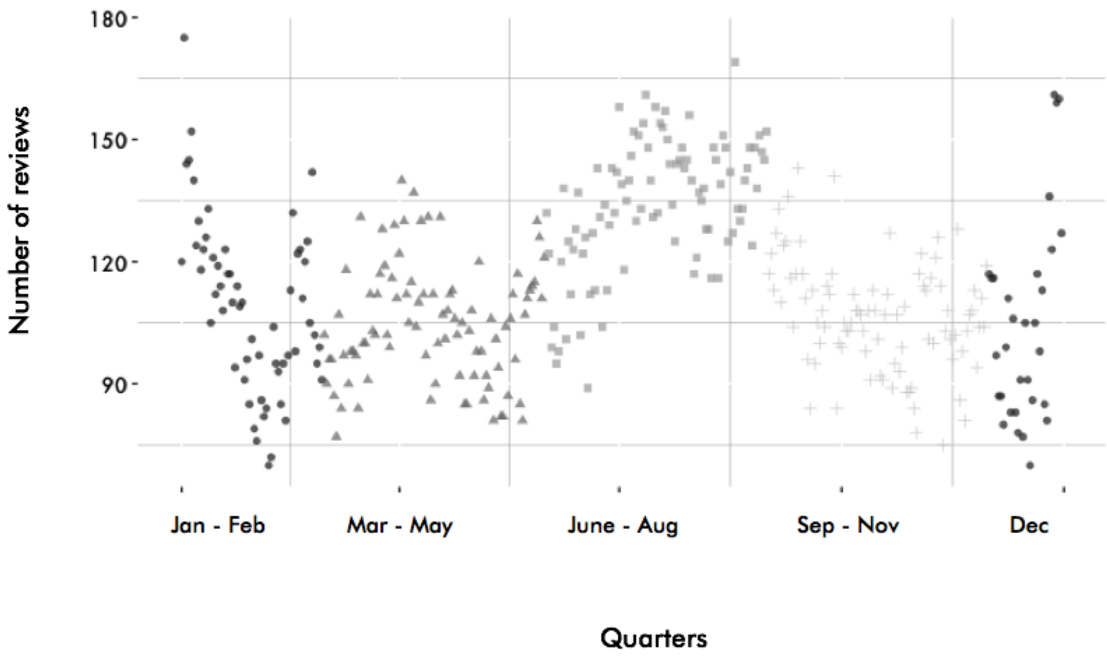


Figure 4. Degree of centrality of terms-network based on correlations.

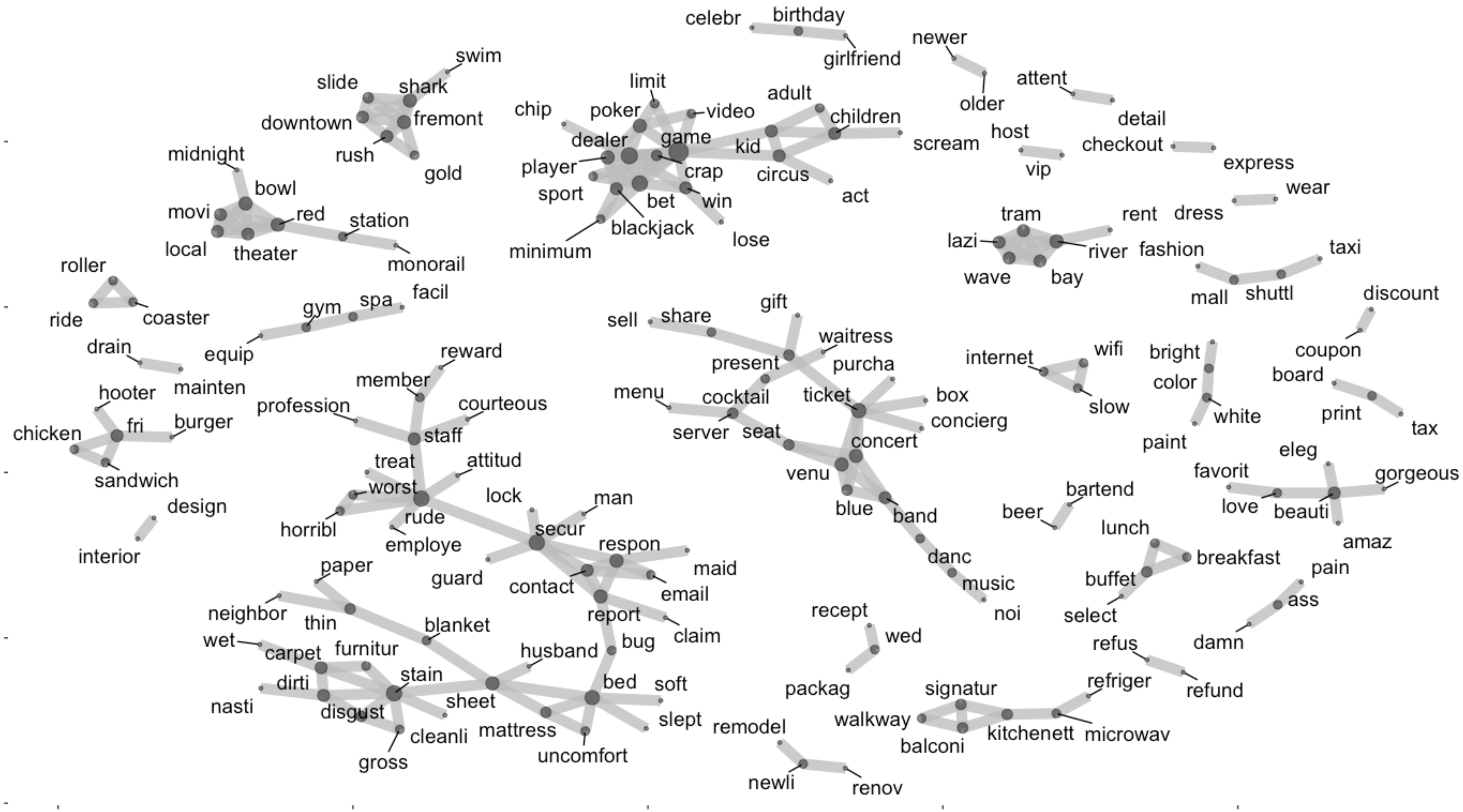


Figure 5. Rate of cross-validated perplexity change.

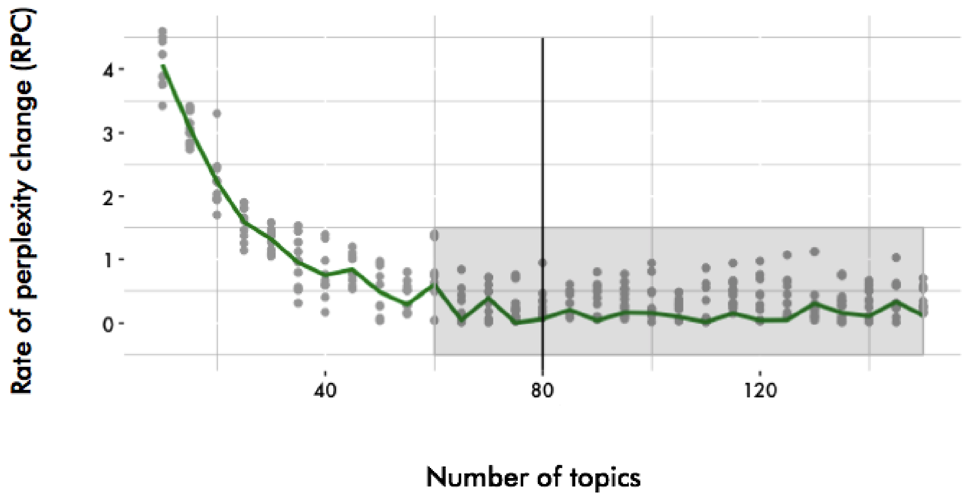


Figure 6a. Top-7 terms to describe topics (λ relevance-value = 0.60).

Figure 6b. Top-7 terms to describe topics (λ relevance-value = 0.00).

Figure 6c. Top-7 terms to describe topics (λ relevance-value = 1.00).

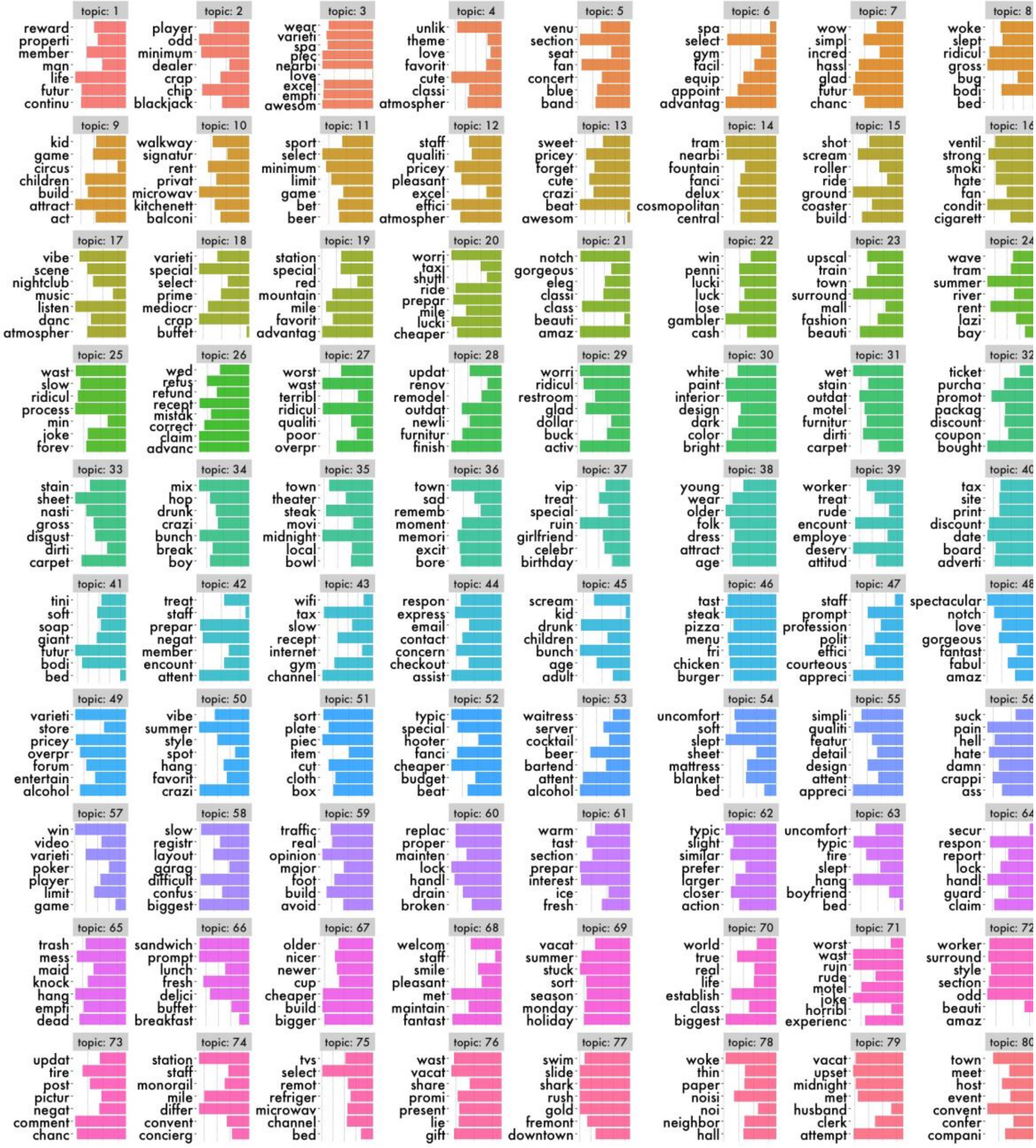
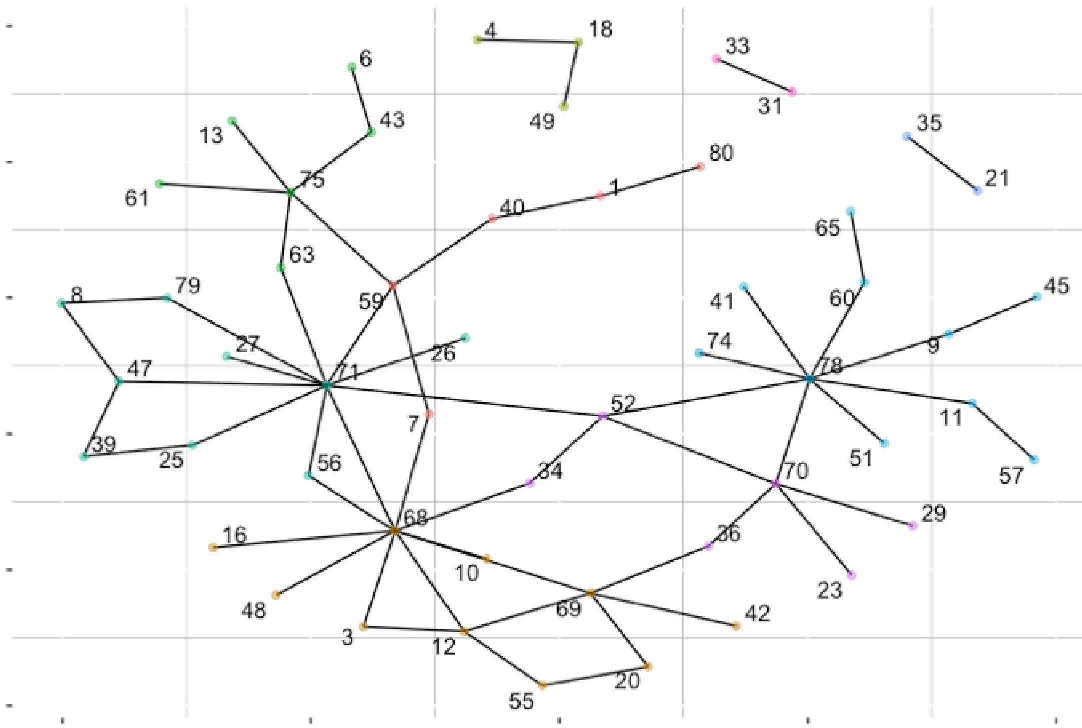


Figure 7. Communities of topics based on modularity index.



Community 1: 1, 40, 59, 80. Community 2: 3, 10, 12, 16, 20, 42, 55, 68, 69. Community 3: 4, 18, 49. Community 4: 6, 13, 43, 61, 63. Community 5: 8, 25, 26, 27, 39, 47, 56, 71, 79. Community 6: 9, 11, 4, 45, 51, 57, 60, 65, 74, 78. Community 7: 21, 35. Community 8: 23, 29, 34, 36, 52, 70. Community 9: 31, 33.

Figure 8. Evolution of star ratings proportion (%) by date (x-axis, from: 2005, to: 2017).

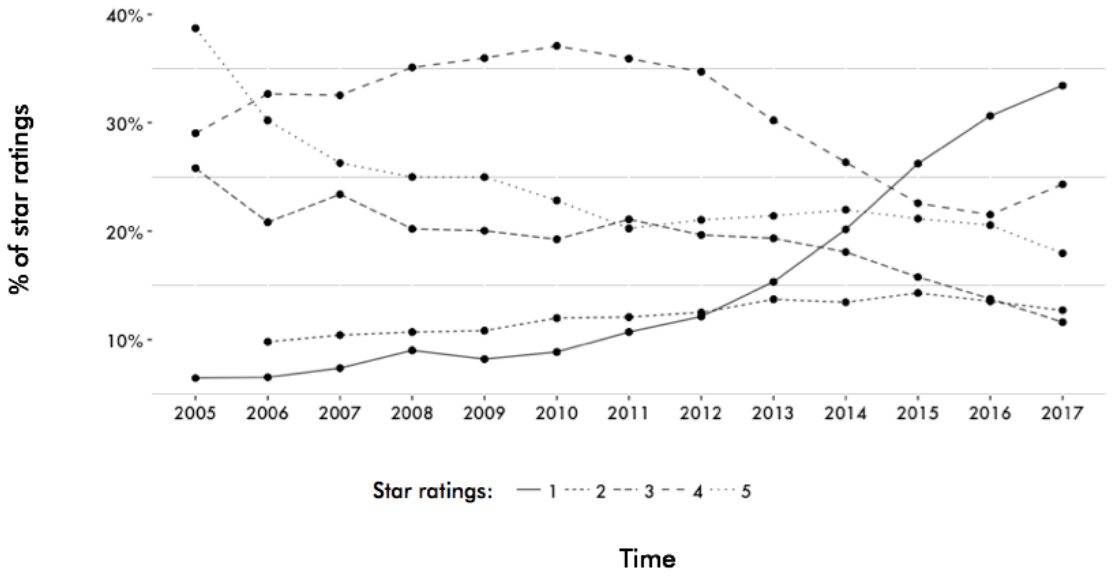


Figure 9. Evolution of averaged (by date) AFINN- and trust scores.

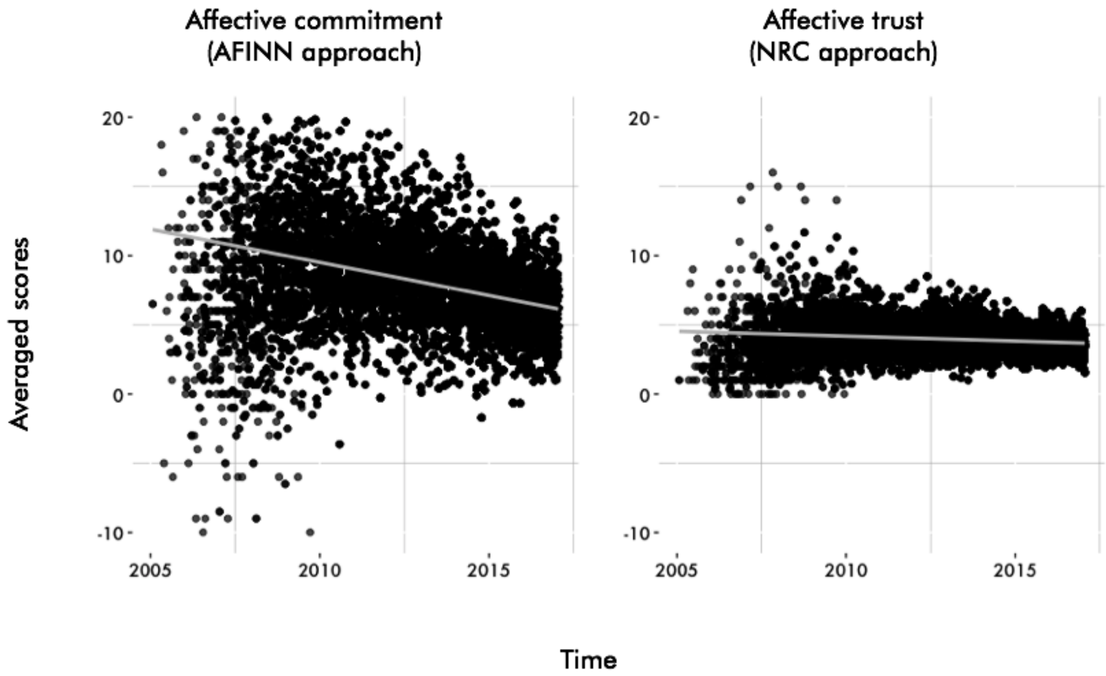


Figure 10. Inner model.

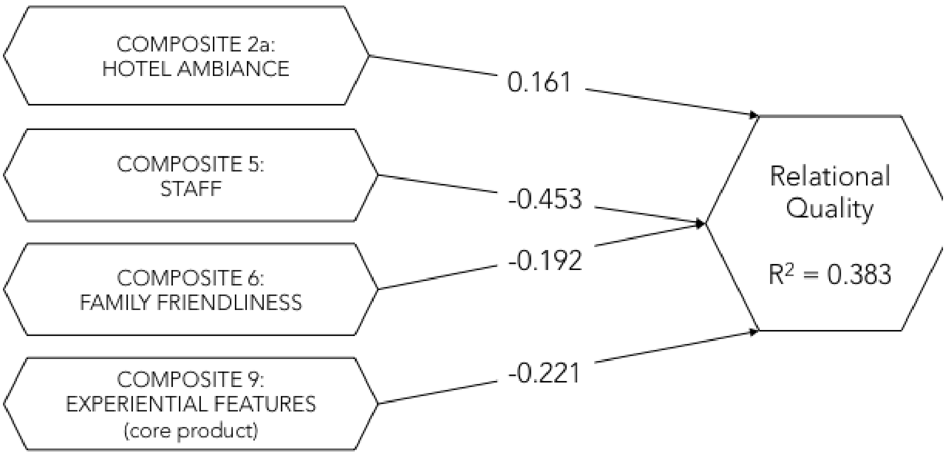


Figure 11. Density plot of out-of-sample (OOS) and in-sample (IS) residuals.

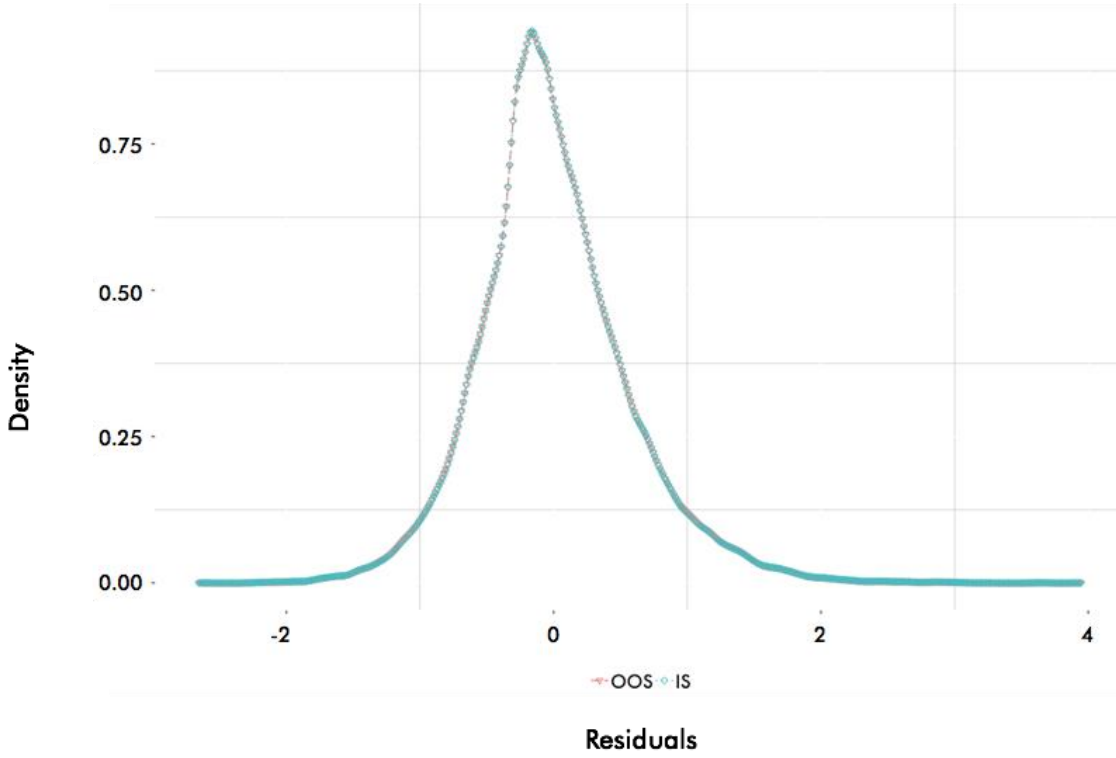


Table 1. Descriptive statistics for RQ variables.

	Mean	SD	Median	Min	Max	Skew	Kurtosis	SE
Rating (Satisfaction)	3.21	1.41	4.00	1.00	5.00	-0.32	-1.21	0.01
NRC-Trust (Trust)	3.92	3.26	3.00	0.00	28.00	1.49	3.16	0.02
AFINN-Sent. (Commitment)	7.83	9.81	7.00	-45.00	80.00	0.47	1.58	0.05

n= 41,413

Table 2. Size and significance of weights.

Composite	Items	Loadings	Weights	VIF	t-statistic	p-value	Confidence Interval	
							Lower 2.5%	Upper 97.5%
Composite 2a: Hotel ambiance	t_03	0.598	0.519	1.012	20.533	0.000	0.476	0.558
	t_10	0.292	0.218	1.009	6.790	0.000	0.164	0.270
	t_12	0.410	0.335	1.012	11.631	0.000	0.287	0.382
	t_20	0.077	0.018	1.006	0.550	0.291	-0.037	0.070
	t_48	0.712	0.635	1.016	27.089	0.000	0.595	0.672
	t_55	0.208	0.141	1.006	4.349	0.000	0.086	0.192
	t_68	0.121	0.044	1.013	1.290	0.099	-0.016	0.097
Composite 2b: Hotel ambiance	t_16	0.684	0.665	1.003	5.888	0.000	0.597	0.726
	t_42	0.341	0.315	1.011	7.948	0.000	0.224	0.403
	t_69	0.671	0.652	1.001	15.786	0.000	0.581	0.716
Composite 5: Staff	t_25	0.309	0.203	1.014	22.270	0.000	0.190	0.226
	t_26	0.362	0.263	1.015	26.456	0.000	0.246	0.279
	t_27	0.496	0.399	1.014	43.733	0.000	0.380	0.416
	t_39	0.415	0.274	1.032	28.174	0.000	0.255	0.293
	t_47	-0.044	-0.084	1.005	8.906	0.000	-0.103	-0.066
	t_56	0.343	0.285	1.006	29.862	0.000	0.266	0.303
	t_71	0.697	0.568	1.036	67.333	0.000	0.551	0.584
	t_79	0.184	0.113	1.009	11.256	0.000	0.093	0.132
	t_08	0.129	0.092	1.004	9.120	0.000	0.072	0.112
Composite 6: Family friendliness	t_11	0.101	0.128	1.016	7.257	0.000	0.094	0.163
	t_41	0.118	0.087	1.004	4.656	0.000	0.049	0.123
	t_45	0.201	0.166	1.023	8.990	0.000	0.130	0.202
	t_51	0.296	0.270	1.003	15.126	0.000	0.234	0.304
	t_57	0.015	0.044	1.020	2.405	0.016	0.009	0.081
	t_60	0.629	0.579	1.007	36.517	0.000	0.548	0.610
	t_65	0.473	0.423	1.004	25.759	0.000	0.390	0.455
	t_74	0.093	0.080	1.002	4.293	0.000	0.042	0.116
	t_78	0.400	0.342	1.007	20.473	0.000	0.308	0.375
	t_09	0.405	0.378	1.024	22.566	0.000	0.344	0.409
Composite 9: Experiential features (core product)	t_31	0.602	0.478	1.025	33.732	0.000	0.450	0.505
	t_33	0.883	0.807	1.025	82.181	0.000	0.787	0.826
RQ: Relational quality	Satisfaction	0.762	0.431	1.698	92.142	0.000	0.423	0.439
	Trust	0.553	0.313	1.330	44.350	0.000	0.303	0.326
	Commitment	0.926	0.538	2.058	235.834	0.000	0.534	0.541

Based on a bootstrapping procedure of 5,000 sub-samples and a two-tail distribution

Table 3. Effect on the endogenous variable.

	Direct effect	t-statistics	p-values	Confidence Interval Lower (2.5%)	Confidence Interval Upper (97.5%)	Explained variance	f^2
RQ (Relational Quality) (R^2= 0.383)							
Composite 2a (Hotel ambiance)	0.161	28.901	0.000	0.152	0.170	2.46%	0.058
Composite 5 (Staff)	-0.453	108.972	0.000	-0.460	-0.446	23.19%	0.320
Composite 6 (Family friendliness)	-0.192	39.679	0.000	-0.201	-0.184	5.53%	0.058
Composite 9 (Experiential features)	-0.221	50.202	0.000	-0.228	-0.214	7.14%	0.076

Based on a two-tail distribution

Table 4. Predictive assessment.

Construct prediction summary	Q^2 (PLS)
RQ	0.237
Indicator prediction summary	**Q^2 (PLS)**
Trust	0.221
Commitment	0.108
Satisfaction	0.337

Note: trust, satisfaction and commitment are indicators of RQ.

About the authors

Manuel J. Sanchez-Franco, PhD is Full Professor of e-Business Management and Social Communication at the Department of Business Administration and Marketing, Universidad de Sevilla, Spain, and collaborates on research projects with companies and public administrations. His research interests are in the areas of consumer behavior and media planning. He has published books and book chapters, as well as research papers in several top ranked journals including Internet Research, Journal of Business Research, Journal of Interactive Marketing, Psychology & Marketing, Computers in Human Behaviour, Information & Management, Electronic Commerce Research and Applications, Computers and Education, Online Information Review, Management Decision, Behaviour and Information Technology, The Service Industries Journal, Physiology & Behavior, among others. Manuel J. Sanchez-Franco is the corresponding author and can be contacted at: majesus@us.es

Gabriel Cepeda-Carrion, PhD is Full Professor in the Management and Marketing Department at the Universidad de Sevilla, Spain. His main research topics include knowledge management, absorptive capacity, dynamic capabilities and organizational learning and unlearning. His research is published in several top ranked journals including the Journal of Knowledge Management, Management Decision, Journal of Business Research, British Journal of Management, Time and Society, Journal of Organizational Change Management, International Journal of Manpower, International Journal of Information Management, Behaviour and Information Technology, The Service Industries Journal, among others. He has also served as Guest Editor of the Service Industries Journal, Journal of Business Research and European Management Journal.

José L. Roldán, PhD is Professor of Management in the Department of Business Administration and Marketing, Universidad de Sevilla, Spain. His current research interests include technology acceptance models, business intelligence, knowledge management, organizational culture, social network sites and partial least squares (PLS). His recent contributions have been published in European Journal of Operational Research, International Journal of Project Management, British Journal of Management, Journal of Business Research, International Business Review, European Journal of Information Systems, International Small Business Journal, Computers in Human Behavior, Industrial Marketing Management, among others. In addition, he has published several book chapters for Idea Group Publishing, Information Science Reference and Springer. He is currently on the editorial board of The Data Base for Advances in Information Systems. He has also served as Guest Editor of the European Journal of Information Systems (EJIS), Journal of Business Research and European Management Journal.