Article

Journal of Vacation Marketing
1–14

DOI: 10.1177/1356766717706103
journals.sagepub.com/home/jvm
SAGE

Determining the reliability and validity of online reputation databases for lodging: Booking.com, TripAdvisor, and HolidayCheck

Manuel Rodríguez Díaz and Tomás F Espino Rodríguez

Universidad de Las Palmas de Gran Canaria, Spain

Abstract

Online reputation is a strategic factor in determining the competitiveness and marketing capacity of lodging companies. The influence of online opinions on customers' decisions is increasing, and, consequently, the online reputation is a new marketing tool to capture clients and reach sales objectives in the lodging industry. In this context, the reliability and validity of customer evaluations available on websites is an essential key to competing in a tourism market influenced by the development of the Internet. The objective of this study is to analyze three of the most important online reputation websites in tourism in order to establish the reliability and validity of the scales used in customer reviews. The results demonstrated that the three websites analyzed fulfill the conventional statistical criteria of reliability and validity. However, a new type of validity is formulated in this study in order to test the capacity of the scales to determine the similarities or differences between tourism goods and services. Nonparametric tests were carried out, demonstrating that although the three websites meet the conventional statistic criteria of reliability and validity, only Booking.com has the capacity to differentiate between destinations.

Keywords

Lodging, online customer review, online reputation, online tourism database, perceived value, service quality

Introduction

Online reputation is currently one of the most important topics in defining hotels' marketing strategy. The number of customers who read the evaluations available on different websites such as Booking.com, TripAdvisor, and HolidayChecks before making a decision about their holiday is increasing (Horster and Gottschalk, 2012). The traditional marketing strategy of lodgings was based on travel agents and tour operator intermediaries specialized in selling hotel beds (Rodríguez-Díaz and Espino-Rodríguez, 2008). However, the Internet has produced a revolution in the relationships between lodgings and guests, breaking with the concentration of sales through commercial channels designed by tour operators.

Corresponding author:
Manuel Rodríguez Díaz, Facultad de Ciencias Económicas, Empresariales y Turismo, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria 30017, Spain.
Email: manuel.rodriguezdiaz@ulpgc.es

The Internet revolution in the tourism sector has had three main effects. First, the Internet provides customers with effective information at a low cost, based on reviews shared by other clients (Yacouel and Fleisher, 2012). Second, the Internet is a useful tool to obtain direct information from customers about the service quality achieved by the lodging, with the possibility of comparing it with main competitors. Third, the influence of the information available on websites is so great that hotel managers need to develop tools and methodologies that show their competitive positioning in the market and facilitate the implementation of tactical decisions (Rodríguez Díaz et al., 2015).

A company's online reputation consists of the set of opinions, experiences, and evaluations of customers shared on websites about a product, service, or brand (Hernández Estártico et al., 2012). From this perspective, the online reputation is out of the scope of companies because it is an external factor. However, customer information obtained on the Internet and the company–customer interrelationships maintained personally or on the website provide companies with a new tool to influence their own online reputation (Gössling et al., 2016).

Online customer reviews have achieved great importance in the implementation of the communication strategy in the tourism industry (Li et al., 2013). The consequence is the appearance of a new dimension in the communication strategy of tourism companies, where websites such as Booking.com, TripAdvisor, and HolidayCheck have developed a new form of communication between clients and lodgings. Different studies identify the impact of online reviews on consumer attitudes, purchase behavior, and services companies' performance (Kim et al., 2015; Lee and Ro, 2016; Vermeulen and Seegers, 2009; Ye et al., 2014), finding them to be a useful measure to control service quality (Chun, 2005; Hernández Estárico et al., 2012). Moreover, price strategies and revenue management take into account the measurement of hotels' quality and perceived value obtained through opinion websites (Varini and Sirsi, 2012; Yacouel and Fleisher, 2012).

In this technological and information-based world, the first basic problem is to determine the level of quality of information available on websites. Therefore, before designing a communication strategy on the web, companies need to correctly answer these questions: (1) Is the information about the online reputation reliable?; (2) Is the information about the online reputation valid?; (3) Can the online reputation be measured with the limited number of variables used by websites?; (4) What is the most reliable and valid online reputation website?; (5) Is it possible to determine the reliability and validity of the online reputation scales through the conventional statistical process alone?; (6) Is it possible to establish whether there is false or tendentious information about the online reputation?

The two objectives of this article are to look for answers to these questions. Specifically, the first objective is to determine the reliability and validity of the scales of variables used by websites, analyzed following the conventional statistical method. Scales that fulfill these criteria of reliability and validity should have the capacity to assimilate or differentiate between goods or services. Therefore, the second objective is to test this capacity by proposing a nonparametric test to measure a new type of validity.

The data available on Booking.com, TripAdvisor, and HolidayCheck are analyzed to establish the reliability and validity of the scales used to measure lodgings' online reputations. Although the number of variables is limited, the study analyzes whether the scales can be tested using the conventional statistical method to determine their reliability and validity. In this context, a new validation test is proposed to show whether the databases studied can set the similarity or discrimination between goods or services in the lodging offer of three tourism destinations.

To achieve this objective, the article begins with a literature review, followed by a description of the methodological characteristics of the research carried out. The results of the Kruskal–Wallis nonparametric test and partial least squares (PLS) analysis, as well as the main implications, are presented in the next section. Finally, we present the conclusions, limitations, and future lines of research.

Literature review

Online customer reviews have great relevance in tourism research because they are a basic resource to evaluate the level of service quality and client satisfaction of lodgings and destinations (Horster and Gottschalk, 2012; Hu et al., 2008; Mudambi and Schuff, 2010; Ye et al., 2014). Information available on websites provides a useful way to know and analyze customers' perceptions (Pantelidis, 2010; Ryu and Han,

2010; Vermeulen and Seegers, 2009; Zhang et al., 2010) and define the competitiveness strategy of lodging companies (Grössling et al., 2016; Rodríguez Díaz et al., 2015; Ye et al., 2014). Online reputation can be defined as the communicative and interactive processes for spreading information exchanged by actors within a social network (Einwiller, 2003).

One aspect of the online reputation available on the web is based on customer evaluations of a limited number of variables because customers are not willing to spend much time giving their opinions. Specialized websites use scales of variables that measure the level of service quality and perceived value of lodgings (Rodríguez Díaz et al., 2015). Value is a subjective concept related to customers' perceptions and attitudes, which have a direct influence on companies' competitiveness (Anderson and Narus, 1998).

Value is considered as a combination of the quality perceived by customers and the price paid for the goods or services acquired (Holbrook, 1994; Oliver, 1997, Parasuraman et al., 1988, Rust and Oliver, 1994; Zeithaml, 1988). In tourism, Prebensen et al. (2012) establish that perceived value is a process through which a tourist shares information based on his/her experiences in a tourism destination or lodging to create a meaningful picture of the value obtained from these experiences.

In the field of tourism, value has been studied in relation to customer satisfaction (Baker and Crompton, 2000; Chadee and Mattsson, 1996; Füller et al., 2006; Li et al., 2013; Nam et al., 2011), the lodging offer of destinations (Tajzadeh-Namin, 2012; Williams and Soutar, 2009), and added value (Jeong, 2002; Rodríguez Díaz et al., 2015). The usual method applied in the literature to measure perceived value consists of a scale of variables (Gallarza and Saura 2006; Sweeney et al., 1999). However, in tourism, due to the small number of variables that can be evaluated online by customers, Prebensen et al. (2012) recognize that perceived value is usually measured by a single item, such as the "quality-price relationship" or "value for money."

Other decisive information available on websites about online reputation consists of the prices and categories of lodgings. Although these variables are independent from direct evaluations by customers, they are necessary in order to evaluate the reliability and validity of the scales used by websites through PLS analysis. We test the predictive validity of the scales used by websites in relation to the value construct. For this reason, we need to include price and category of lodgings variables in the analysis, along with service quality variables. The category is related to perceived value because a higher category should obtain a higher price for services delivered (Abrate et al., 2011; Israeli, 2002; López Fernández and Serrano Bedia, 2004; Núñez-Serrano et al., 2014; Tanford et al., 2012).

Yacouel and Fleischer (2012) establish that online trustworthiness and credibility are crucial in developing the online reputation of lodgings without false or tendentious opinions. Gössling et al. (2016:4) consider that "trust and credibility are key aspects of online evaluations, because trust, defined [. . .] as the belief that online content is reliable, and credibility, that is, the condition of being considered honest, are closely related to consumer choices."

The key methodological aspects used in this research are the reliability and validity of attitudes measurement scales. Hair et al. (1999) establish that reliability determines the degree of consistency between multiple measurements of a variable. Hayes (1992) defines reliability as the degree to which measurements are free from deviations produced by causal errors. Bagozzi (1996) considers that there are two primary types of reliability: (1) test–retest reliability or stability, which establishes the consistency between the same measurements repeated at different moments in time; and (2) internal consistency, which establishes the degree of agreement between two or more measurements of the same theoretical concept, where Cronbach's α and composite reliability (CR) (Fornell and Lacker, 1981) are the most commonly used methods.

Once the scale fits a conceptual definition and reaches the required levels of reliability, the researcher must determine the scale's validity. Validation is the extent to which a scale or a set of measures accurately represents the concept studied (Hair et al., 1999). The three most widely accepted forms are convergent, discriminant, and predictive validity. Convergent validity values the degree to which two measures of the same concept are correlated. High correlations indicate that the scale measures the studied concept (Hair et al., 1999). Convergent validity is tested in this study through the average variance extracted (AVE) index calculated using PLS analysis. Discriminant validity is the degree to which two conceptually similar concepts differ from each other. In this case, the correlation between concepts should be low (Hayes, 1992).

Discriminant validity is determined by the correlations matrix between the concepts analyzed. Finally, predictive validity reflects the degree to which a scale is capable of making accurate predictions of other concepts that are theoretically grounded in prior research. Bagozzi (1996) states that predictive validity is very similar to nomological validity, but predictive validity is a bivariate relationship, whereas nomological validity is a multivariate relationship based on a previous theoretical model. Predictive validity is tested in this study through PLS analysis, where the independent variables are the service quality scales, prices, and categories of lodgings (number of stars), whereas the dependent variable is the perceived value.

If a scale fulfills the tests of reliability and validity, the measurement of two or more similar goods or services should produce similar results. By contrast, if the scale is used to measure two or more different goods or services, the results should be discriminant. The new validity proposed is focused on contrasting this circumstance, and it is called validity of similarities or differences between goods or services. A nonparametric test is carried out where the dependent variable is the three tourism destinations and the independent variables are the service quality scale, the price, and the category of lodgings (number of stars). The results demonstrated that only the scale used by Booking.com fulfills this new validity.

Research methodology

To carry out the study, information was collected about the lodging offer of three tourism destinations (Gran Canaria and Tenerife in Spain and Agadir in Morocco) from the websites of Booking.com, TripAdvisor, and HolidayCheck. These destinations are located very close in the Atlantic Ocean, with a distance between the Canary Islands and Agadir of 420 km. This is crucial because the study objective is to determine the discrimination capacity of the website. Therefore, the three destinations have to compete with each other in order to achieve this aim.

In Gran Canaria, 149 lodgings were analyzed in the southern part of the island, whereas in the southern part of Tenerife, 49 tourism complexes were studied. In addition, in Agadir, which is located in southern Morocco, 38 lodgings were analyzed. The hotels were selected based on the condition that the information had to be available in the three online databases studied. Of the total number of customer comments collected in the databases analyzed, 59.895 were from Booking.com, 104.500 from TripAdvisor, and 51.725 from HolidayCheck.

Two groups of variables were used in the study. The first group includes variables that measure the objective and structural characteristics of lodgings, such as the category, based on the number of stars (one to five stars), and the price. A peculiarity of lodging prices is that they change with the season. For this reason, in the statistical analysis, we utilized two prices: (1) the most common price in the high season and (2) the most common price in the low season. As these variables are considered constant, they were used to evaluate the three databases analyzed.

Each online database used a specific scale to obtain information from customers. In some cases, they used similar variables, such as location, but normally there are significant differences in their scales. Another important difference is the type of scale employed, as Booking.com obtains information from customers using a 10-point scale (1 = *very poor evaluation*; 10 = *excellent*), whereas TripAdvisor uses a five-point scale (1 = *very poor evaluation*; 5 = *excellent*), and HolidayCheck uses a six-point scale (1 = *very poor evaluation*; 6 = *excellent*).

The variables evaluated by Booking.com are staff, comfort, facilities, location, cleanliness, Wi-Fi, and value for money. Because the Wi-Fi system depends on the level of development of the public and private infrastructures of the destinations, we did not include it in the study. On the other hand, TripAdvisor uses the variables location, sleep quality, comfort, facilities, cleanliness, and quality/price relationship. According to Ye et al. (2014) and Rodríguez Díaz et al. (2015), the scales used by Booking.com and TripAdvisor can be divided into two parts: (1) the variables measuring the service quality (staff, comfort, facilities, location, cleanliness, sleep quality) and (2) the variables measuring the perceived value of the lodging (value for money or quality/price relationship).

By contrast, HolidayCheck assesses the variables room, service, location, gastronomy, and sport and leisure. As all these variables measure different characteristics of service quality, and this website does not have any variable that measures perceived value, we decided to use the percentage of customer recommendations as the variable "value perceived by customers" in the PLS analysis. The gastronomy variable was not

included in the statistical analysis because it is only evaluated in HolidayCheck, and many lodgings do not offer restaurant service.

To determine the reliability and validity of the scales used by websites with customer tourism reviews, the methodology applied is the following.

1. Carry out the Kruskal–Wallis nonparametric test to establish the validity of the similarity or difference between goods or services. The objective is to establish whether the scales used by the websites have the capacity to detect similarities between goods or services that are quite similar or dissimilarities if the goods or services are different. The dependent variable utilized in the test consists of the three tourism destinations studied, which represent the services compared, whereas the independent variables are the service quality scales, prices, and categories of lodgings (number of stars).
2. Determine the scales' reliability with Cronbach's α and CR (Fornell and Lacker, 1981).
3. Establish the scales' convergent, discriminant, and predictive validity (Fornell and Lacker, 1981; O'Cass and Ngo, 2007; Roldán and Sánchez-Franco, 2012). To test convergent validity, the AVE index was calculated. The correlation matrix was used to determine discriminant validity, whereas the structural analysis PLS was applied to test predictive validity (Chin, 1998; Lohmöller, 1989). Following Bagozzy (1996), predictive validity is shown when a construct has the capacity to predict another variable or construct that is theoretically different but closely related. This is the case of the concept of value, which in the academic literature is considered to be directly related to service quality (Rust and Oliver, 1994). The use of PLS analysis in this study is justified by the minor assumptions required in relation to the covariance methods, such as Lisrel, EQS, or Amos (Chin et al., 2003). Likewise, because the test is not complex and the analyzed scales have few variables, PLS is more appropriate to guarantee the robustness of the results (Frías-Jamilena et al., 2012; Qureshi and Compeau, 2009).

Table 1. Kruskal–Wallis test of structural variables.

Variables	Sig.
Category (number of stars)	0.000
Price in high season	0.000
Price in low season	0.020

Table 2. Kruskal–Wallis test of Booking.com.

Variables	Sig.
Cleanliness	0.065
Comfort	0.286
Location	0.007
Facilities	0.079
Staff	0.010
Value for money (quality–price relationships)	0.000

Analysis of results

Nonparametric analysis

Four Kruskal–Wallis nonparametric tests were performed, using the tourism destinations as dependent variable. The tourism destinations are used to represent the services to be compared to determine the validity of similarities or differences of goods and services. The first analysis was carried out to establish whether the structural variables (category or number of stars, price in high season, and price in low season) are related to the destinations.

Table 1 shows that the category (number of stars) and price in high season are significant ($p < 0.001$), whereas the price in low season is also significant at 5% ($p < 0.020$). These results demonstrate that the composition of the lodging offer is different, and the price in high season oscillates depending on the level of competitiveness of the destinations studied. Moreover, the price in low season tends to go down and be more similar across destinations, perhaps due to hotels' special promotions. However, the results also show significant price differences in the low season.

The results of the Kruskal–Wallis nonparametric test applied to Booking.com show that the variables location, staff, and value for money are significant at 5%, whereas cleanliness and facilities are significant at 10% (see Table 2). The only non-significant variable is comfort. These results reveal that the scale of variables used by Booking.com has the capacity to discriminate between destinations on most of the variables. The value for money (quality/price relationships)

Table 3. Kruskal–Wallis test of TripAdvisor.

Variables	Sig.
Location	0.150
Sleep quality	0.710
Room	0.240
Service	0.262
Cleanliness	0.708
Quality–price relationships	0.266

Table 4. Kruskal–Wallis test of HolidayCheck.

Variables	Sig.
Percentage of recommendation	0.491
Room	0.575
Service	0.536
Location	0.000
Sports and entertainment	0.592

variable is the most significant ($p < 0.001$), followed by location ($p = 0.007$) and staff ($p = 0.010$). Likewise, cleanliness ($p = 0.065$) and facilities ($p = 0.079$) also show differentiations between destinations, but at a low level.

Table 3 shows that none of the variables used by TripAdvisor is significant. Therefore, they do not have the capacity to differentiate between destinations. Moreover, the results obtained for HolidayCheck are similar, with the exception of the variable location ($p < 0.001$). Table 4 shows the nonparametric test carried out on the variables from HolidayCheck, demonstrating that the scale used does not differentiate between tourism destinations. In conclusion, the variables used by TripAdvisor and HolidayCheck do not fulfill the validity of similarities and differentiations between goods and service, whereas Booking.com is the only database that fulfills this validity for most of the scale's variables.

PLS analysis

Using PLS analysis (SmartPLS 2.0 M3), developed by Ringle et al. (1995), it is possible to determine the reliability and validity of the online scales used by Booking.com, TripAdvisor, and HolidayCheck. PLS analysis is carried out in two stages. First, the reliability and validity of the measurement model is evaluated (Chin, 1998; Roldán and Sánchez-Franco, 2012). Second, the structural model is examined to test the predictive validity of the scales (Barclay et al., 1995; Falk and Miller, 1992).

Measurement model. Table 5 shows the factor loadings of the variables. All the variables exceed the accepted level of 0.707 (Carmines and Zeller, 1979), with the exception of location, where only Booking.com exceeds the level of 0.50 (Falk and Miller, 1992; Hasan and Ali, 2007). It can be observed that the variable location has low scores in all the databases, as the best offer in terms of quality does not always have the best location.

Reliability is tested by Cronbach's α and CR (Fornell and Lacker, 1981). All the service quality scales obtained a Cronbach's α above 0.7, as did the two price variables (Nunnally, 1978). Convergent validity is tested through the AVE index, where all the service quality and price scales exceeded the value of 0.5. Therefore, the results validate the reliability and convergent validity of the variables used by the online reputation databases for lodging.

Table 6 shows the correlation coefficients, taking into account that the scores on the main diagonal are established by means of the square root of the AVE values. Table 6 reveals that the values on the main diagonal are higher than the correlations indexes of the other variables (Fornell and Lacker, 1981; O'Cass and Ngo, 2007; Roldán and Sánchez-Franco, 2012). These results confirm that the variables used by Booking.com, TripAdvisor, and HolidayCheck to measure service quality and perceived value also have discriminant validity. It is important to highlight that the highest correlation index is between service quality and value, where the coefficients exceed the score of 0.70. In conclusion, the study demonstrates that the three databases observe the common statistical criteria for reliability and convergent and discriminant validity.

Structural model. The model formulated in the PLS analysis was carried out in the three databases (see Figures 1 to 3). The variables used to evaluate lodgings' prices and categories were the same in the three models because they are structural variables. However, the scales utilized to measure service quality are different for some variables. Moreover, perceived value was measured through a direct variable on Booking.com (quality for money) and TripAdvisor (quality–price relationship), whereas for HolidayCheck we used the percentage of recommendations because this database does not have a specific variable available.

Table 5. Evaluation of the measurement model.

Models, factors, and variables	Factor loadings	t	Composite reliability	AVE	Cronbach's α
Booking.com					
Service quality			0.934	0.745	0.906
Comfort	0.936	127.225			
Facilities	0.952	189.276			
Cleanliness	0.919	101.106			
Staff	0.904	67.145			
Location	0.531	9.945			
Price			0.993	0.986	0.987
Price in high season	0.990	12.763			
Price in low season	0.996	15.573			
Category			1	1	1
Category (no of stars)	1				
TripAdvisor					
Service quality			0.903	0.662	0.863
Sleep quality	0.836	21.317			
Room	0.901	42.524			
Cleanliness	0.868	28.503			
Service	0.907	48.440			
Location	0.471	5.804			
Price			0.993	0.987	0.987
Price in high season	0.994	341.240			
Price in low season	0.992	130.144			
Category			1	1	1
Category (no of stars)	1				
HolidayCheck					
Service quality			0.842	0.586	0.753
Room	0.876	44.352			
Sports and entertainment	0.789	23.123			
Service	0.878	49.805			
Location	0.428	4.278			
Price			0.994	0.988	0.988
Price in high season	0.995	43.855			
Price in low season	0.992	61.593			
Category			1	1	1
Category (no of stars)	1				

Results of the structural models are shown in Table 7. The R squared (R^2) for the three models are greater than 0.5, demonstrating the predictive validity of the scales used by the online databases analyzed. The greatest R^2 was obtained by Booking.com (0.703), followed by HolidayCheck (0.621) and TripAdvisor (0.510). Likewise, goodness-of-fit was also calculated (Tenenhaus et al., 2005), with the three models obtaining scores much higher than 0.36 (Chin, 1998).

Examining the results obtained in Table 7, Booking.com obtained a β score of 0.880, with a significant t of 27.381 in the service quality factor. According to the theoretical proposals, price ($\beta = -0.166$; $t = 3.943$) and category ($\beta = -0.254$; $t = 4.668$) are related negatively and significantly to perceived value. HolidayCheck also achieved useful results because the β coefficient for service quality in predicting the perceived value concept was 0.745, with a t-value of 10.388. The factors price and category maintain inverse and significant relationships with perceived value. Finally, TripAdvisor achieves slightly worse results. Although service quality and category obtained significant scores in relation to perceived value, the price obtained a β coefficient of -0.044, with a non-significant t-value of 1.405. These results demonstrate the predictive validity of the scales used by the three websites, with the exception of the price variable in TripAdvisor.

Discussion of results

Today, the online reputation of lodgings affects their ability to compete. Communication and

Table 6. Latent variables' correlations (AVE square root on diagonal).

	Quality	Price	Category	Perceived value
Booking.com				
Service quality	0.996			
Price	0.259	0.961		
Category (no of stars)	0.276	0.413	1	
Perceived value	0.766	−0.079	−0.043	1
TripAdvisor				
Service Quality	0.950			
Price	0.260	0.996		
Category (no of stars)	0.314	0.411	1	
Perceived value	0.706	0.114	0.129	1
HolidayCheck				
Service quality	0.917			
Price	0.262	0.996		
Category (no of stars)	0.328	0.413	1	
Perceived value (percentage recommendation)	0.756	0.076	0.088	1

Note: The scores on the diagonal are the square root of the AVE, and the other scores are the correlation coefficients between latent variables.

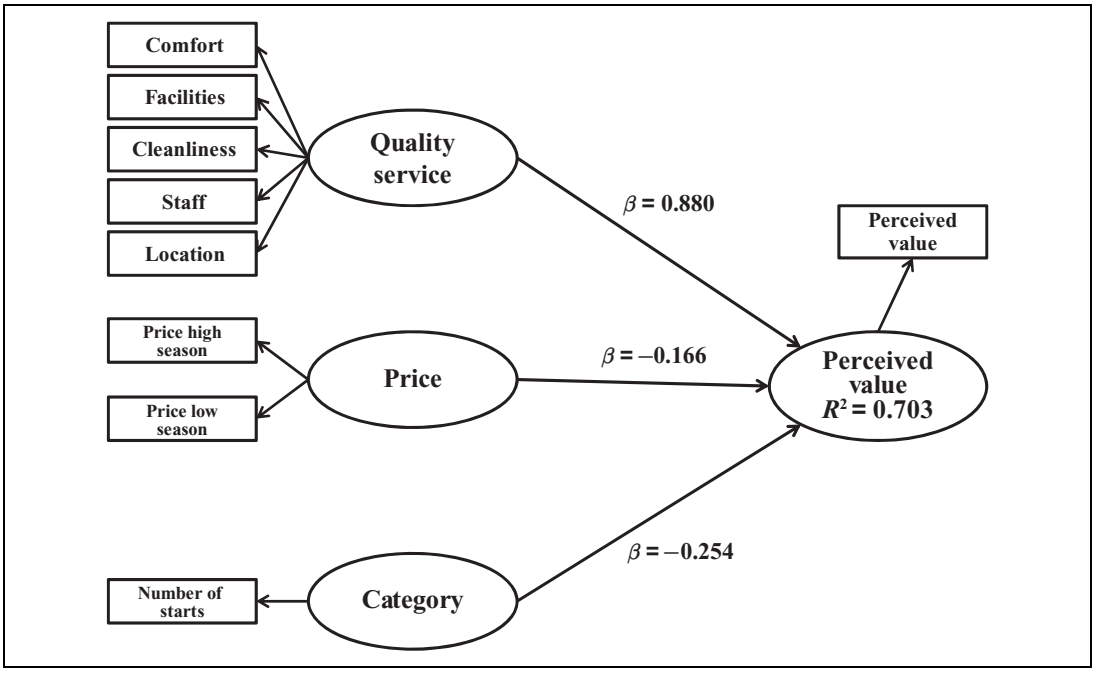

Figure 1. Structural model of variables used by Booking.com.

price strategies are based on customer perceptions displayed on specialized websites such as Booking.com, TripAdvisor, and HolidayCheck. The proposed methodology was tested to determine whether the scales and measurements used by the three databases studied are reliable and valid, and which is the most reliable and valid. The results obtained show that the three websites are reliable and valid using the traditional statistic procedures. That is, they meet the requirements of reliability and convergent, discriminant, and predictive validity.

The main characteristic of the scales used by the websites is that they have few variables because it is very difficult to obtain much more information from customers in an online survey. Another characteristic is that the variables used by the websites studied do not always coincide. Taking these circumstances into account, the statistical analysis carried out demonstrated that all the scales used to measure service quality and price are reliable because they obtained a Cronbach's α greater than 0.7, and the AVE indexes exceeded 0.5.

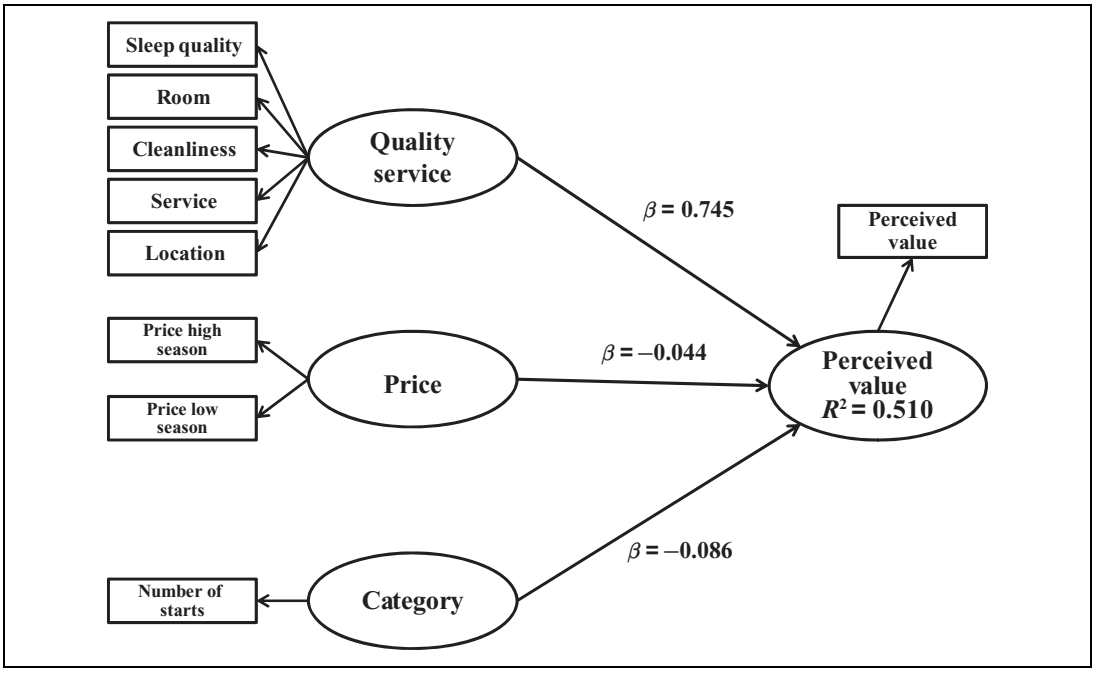

Figure 2. Structural model of variables used by TripAdvisor.

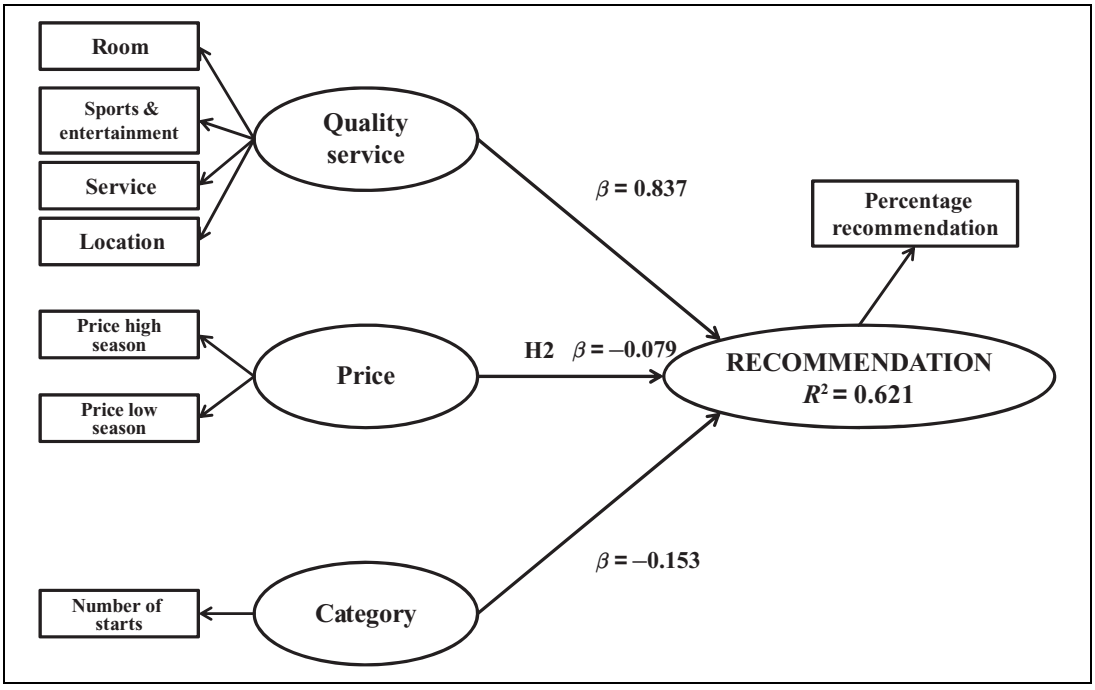

Figure 3. Structural model of variables used by HolidayCheck.

The study also demonstrated that all the scales have convergent, discriminant, and predictive validity, except the price factor on TripAdvisor, which is not significant in predicting the value perceived by customers. Although HolidayCheck does not have a specific variable to measure the perceived value concept, in contrast to Booking.com (value for money) and TripAdvisor (quality–price relationships), the R^2 obtained in the PLS analysis showed that the percentage of customer recommendations is a similar variable to perceived value.

Therefore, is it possible to conclude that the three scales are useful to determine the online

Table 7. Results of the structural models.

Relations	Coefficient	*t*-Value	R^2	Q^2	GoF
Booking.com					
Service quality → *p* value	0.880	27.381			
Price → perceived value	−0.166	3.943	0.703	0.640	0.809
Category → *p* value	−0.254	4.668			
TripAdvisor					
Service quality → *p* value	0.745	10.388			
Price → perceived value	−0.044	1.405	0.510	0.447	0.682
Category → *p* value	−0.086	2.288			
HolidayCheck					
Service quality → percentage recommendation	0.837	24.949			
Price → percentage recommendation	−0.079	1.854	0.621	0.591	0.745
Category → percentage recommendation	−0.153	3.123			

reputation of lodgings? If we follow the traditional statistical methodology to test the scales' reliability and validity, the answer has to be yes. Furthermore, based on the results, it is possible to conclude that the best scores were obtained by Booking.com, followed by HolidayCheck and TripAdvisor. However, the Kruskal–Wallis nonparametric tests carried out demonstrated that only Booking.com can differentiate between tourism destinations on the majority of the variables on its scale, whereas the scales used by TripAdvisor and HolidayCheck do not have this differentiation capacity.

This result confirms the need to include the new validity formulated in this study regarding similarities or differences between goods and services. Normally, in practice, it is very difficult to test scales from this perspective because the limitations of empirical studies make it difficult to obtain data to analyze the similarities and differences between specific goods or services. However, in this study, the information available in the three databases allows us to analyze the capacity of differentiation or similarity between goods or services offered by lodgings in tourism destinations.

The implications of these discoveries are relevant for websites, practitioners, and customers. Websites need to confirm the reliability and validity of the information available about lodgings. However, there are relevant doubts about the possibility of including false and tendentious information in order to improve or, in some cases, harm the online reputation of specific lodgings. In this context, to guarantee their reputation, websites need to develop systems to detect false opinions.

The methodology proposed in this study provides a way to compare the reliability and validity of the scales and opinions available on the websites. If websites do not fulfill these requirements, their reliability can be questioned. Furthermore, the Kruskal–Wallis nonparametric test showed that only Booking.com fulfills the validity of similarities or differences between goods or services.

Tourism websites should provide scientific verification without reducing the number of variables the scales use. They must also try to improve the scales by including variables that are similar to competitor websites, along with a key concept related to perceived value. If a website reduces or simplifies its variables, it will miss the opportunity for verification using scientific methods, thus decreasing its reliability.

From the perspective of practitioners, knowledge about the best method to determine scales' reliability and validity is necessary in order to evaluate market information. First, the variables used to establish the online reputation provide an up-to-date evaluation of the quality of service delivered by lodgings. In this context, it is critical to establish the best source in order to maintain control over processes and activities. Second, the analysis of the competitiveness and strategic positioning of a company and the main competitors also requires reliable information. Finally, practitioners must improve the online reputation of their lodgings and obtain the most reliable and valid information to manage the company.

In the case of customers, when planning their holidays, they need sources of real, updated, and reliable data based on scientific criteria. A website that cannot differentiate between distinct tourism destinations because it provides similar information is not useful for making decisions. Websites offer a wide variety of types of information about the tourism sector, destinations,

and lodgings, and customers can get lost in a sea of data. Therefore, to detect reliable information to evaluate the alternatives for holiday trips, it is important to validate the scales based on their capacity to determine similarities or differences between goods or services.

Conclusions

Online reputation is increasingly important in lodgings' competitiveness and sales strategy. Key issues such as revenue management, image, and communication are based on customer online evaluations on the main websites. However, there are no studies that establish the reliability and validity of the information available on tourism websites such as Booking.com, TripAdvisor, and HolidayCheck. The main problem in achieving the objectives of the article is the limited number of variables used by the websites to measure the online opinions of customers.

The study highlights the need to use scales with a minimum number of variables in order to determine their statistical reliability and validity. Likewise, it is necessary to use a scale to measure service quality, and at least one variable to measure the perceived value, which is critical in testing predictive validity. Reliability was measured through Cronbach's α and CR, obtaining useful scores in all three cases. In relation to convergent and discriminant validity, the results were also successful on the websites. Moreover, predictive validity was tested on Booking.com and HolidayCheck, whereas on TripAdvisor, only the relationship between price and perceived value was not significant. In all three databases, the relationships between service quality and the category with regard to the variable "value perceived by customers" were validated.

Based on the results presented above, we can draw three main conclusions. First, online reputation websites should not reduce the number of variables because they can miss the opportunity to reinforce their image through scientific recognition of the reliability and validity of the available information. Furthermore, a website that reduces variables or uses only one global evaluation of lodgings cannot be tested; therefore, it is not possible to establish the reliability and validity of the information. Likewise, less information available on websites means that customers will be less able to make useful decisions.

Second, it is very important to measure perceived value by customers in order to test predictive validity. In the study carried out, the website HolidayCheck did not evaluate perceived value; consequently, predictive validity was tested by using another similar variable, the percentage of customer recommendations. In this case, it would be advisable in the future to include perceived value as a new variable in order to unify the most important concepts on the scales used in different lodging reputation online databases. However, results confirm that the percentage of recommendations is a useful index to validate predictive validity. Third, the study has demonstrated that the three scales of the websites analyzed are reliable and valid based on conventional statistical procedures defined in the academic literature. A comparison of the results allows us to state that Booking.com obtained slightly better results than the other databases studied.

Nevertheless, the Kruskal–Wallis nonparametric test demonstrated that only the Booking.com database can differentiate between goods and services. This new test to validate the measurement scales of perceptions and attitudes in tourism reveals the need to confirm the validity of scales based on the similarity or differentiation of the goods or services evaluated by customers. This study verifies that TripAdvisor and HolidayCheck do not differentiate between the destinations analyzed; therefore, the information provided to customers does not help them to make their decisions because they cannot discriminate between the lodging offers in the tourism destinations. In other words, TripAdvisor and HolidayCheck provide very similar information about lodgings in different tourism destinations with different lodging offers.

In sum, scales used on websites to measure the online reputation according to customers must also test the validity of the similarities or differences between goods or services. This new contribution can guide future studies that propose new methods to determine the differentiation between goods and services and the processes to test this new validity. In the tourism sector, and taking into account the online information available, this study shows that it is possible to test similarities or differences between the lodgings of three tourism destinations. However, other future studies should test this new validity in specific hotels addressed individually or in clusters by categories.

The implications of this study can be analyzed from the points of view of practitioners and researchers. There are two types of practitioners

involved. On the one hand, for managers of online databases, this study can provide a practical methodology to test and improve the scales used. The minimum number of variables, the perceived value concept, and the procedure to determine the reliability and validity of the scales used to establish the online reputation of lodgings are essential in evaluating the level of trustworthiness of the information provided to customers. One of the websites' main problems is checking authentic comments and removing suspicious or false opinions. From the point of view of lodging managers, knowledge about the reliability and validity of the information collected in online databases is crucial to analyzing competitive positioning in the market and implementing marketing and price strategies.

Finally, this study has implications for research because it represents a new line of research to evaluate online reputation and propose new models and methodologies to determine the reliability and validity of the scales used. In addition, it is very important to establish the main variables to measure service quality and perceived value, as well as the minimum number of variables to include in scales, in order to guarantee the truthfulness of the information available on the websites. Although the main limitation of the study is the low number of variables available on the web, the results show that the scales used by websites are useful for testing their reliability and validity.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Abrate G, Capriello A and Fraquelli G (2011) When quality signals talk: evidence from the Turin hotel industry. *Tourism Management* 32: 912–921.

Anderson JC and Narus JA (1998) Business marketing understand what customer value. *Harvard Business Review* 76(6): 53–65.

Bagozzy RP (1996) Measurement in marketing research: basic principles of questionnaire design. In: Bagozzy RP (ed) *Principles of Marketing Research*. Cambridge: Blackwell Publishers, pp. 1–49.

Baker DA and Crompton JL (2000) Quality, satisfaction and behavioural intentions. *Annals of Tourism Research* 27(3): 785–804.

Barclay D, Higgings C and Thompson R (1995) The partial least squares (PLS) approach to casual modeling: personal computer adoption and use as illustration. *Technological Studies, Special Issue on Research Methodology* 2(2): 285–309.

Carmines E and Zeller R (1979) *Reliability and Validity Assessment*. Newbury Park: Sage Publications.

Chadee DD and Mattsson J (1996) An empirical assessment of customer satisfaction in tourism. *The Service Industries Journal* 16(3): 305–320.

Chin W (1998) The partial least squares approach for structural equation modeling. In: Marcoulides GA (ed) *Modern Methods for Business Research*. London: Lawrence Erlbaum Associates, pp. 295–236.

Chin WW, Marcolin BL and Newsted PR (2003) A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research* 14(2): 189–217.

Chun R (2005) Corporate reputation: meaning and measurement. *International Journal of Management Review* 7(2): 91–109.

Einwiller S (2003) *Vertrauen Durch reputation im elektronishech handel*. PhD Thesis, University of St. Gallen, Switzerland.

Falk R and Miller N (1992) *A Primer for Soft Modeling*. Akron: The University of Akron Press.

Fornell C and Larcker D (1981) Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1): 39–50.

Frías-Jamilena DM, Del Barrio-García S and López-Moreno L (2012) Determinants of satisfaction with holidays and hospitality in rural tourism in Spain: the moderating effect of tourists' previous experience. *Cornell Hospitality Quarterly* 54(3): 294–307.

Füller J, Matzler K and Faullant R (2006) Asymmetric effects in customer satisfaction. *Annals of Tourism Research* 33(4): 1159–1163.

Gallarza MG and Saura IG (2006) Value dimensions, perceived value, satisfaction and loyalty: an investigation of university students' travel behaviour. *Tourism Management* 27(3): 437–452.

Gössling S, Hall CM and Anderson AC (2016) The manager's dilemma: a conceptualization of online review manipulation strategies. *Current Issues in Tourism*. Available at: http://www.tandfonline.com/doi/full/10.1080/13683500.2015.1127337 (accessed 12 February 2016).

Hair JF Jr, Anderson RE, Tatham RL, et al. (1999) *Multivariate Data Analysis*. Upper Saddle River: Prentice Hall International, Inc.

Hasan B and Ali J (2007) An empirical examination of factors affecting group effectiveness in information systems projects. *Decision Sciences Journal of Innovative Education* 5(2): 229–243.

Hayes BE (1992) *Measuring Customer Satisfaction*. Milwaukee: ASQC Quality Press.

Hernández Estárico E, Fuentes Medina M and Morini Marrero S (2012) Una aproximación a la reputación en línea de los establecimientos hoteleros españoles. *Papers de Turisme* 52: 63–88.

Holbrook MB (1994) The nature of customer value: an axiology of services in the consumption experience. In: Rust RT and Oliver RL (eds) *Service Quality: New Directions in Theory and Practice*. Thousand Oaks: Sage Publications, pp. 21–71.

Horster E and Gottschalk C (2012) Computer-assisted webnography: a new approach to online reputation management in tourism. *Journal of Vacation Marketing* 18(3): 229–238.

Hu N, Liu L and Zhang JJ (2008) Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management* 9: 201–214.

Israeli AA (2002) Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. *International Journal of Hospitality Management* 21: 405–424.

Jeong M (2002) Evaluating value-added lodging web sites from customers' perspectives. *International Journal of Hospitality & Tourism Administration* 3(1): 49–60.

Kim WG, Lim H and Brymer RA (2015) The effectiveness of managing social media on hotel performance. *International Journal of Hospitality Management* 44: 165–171.

Lee SH and Ro H (2016) The impact of online reviews on attitude changes: the differential effects of review attributes and consumer knowledge. *International Journal of Hospitality Management* 56: 1–9.

Li H, Ye Q and Law R (2013) Determinants of customer satisfaction in the hotel industry: an application of online review analysis. *Asia Pacific Journal of Tourism Research* 18(7): 784–802.

Lohmoeller JB (1989) *Latent Variable Path Analysis With Partial Least Squares*. New York: Springer-Verlag.

López Fernández MC and Serrano Bedia AM (2004) Is the hotel classification system a good indicator of hotel quality? An application in Spain. *Tourism Management* 25: 771–775.

Mudambi SM and Schuff D (2010) What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly* 34(1): 185–200.

Nam J, Ekinci Y and Whyatt G (2011) Brand equity, brand loyalty and consumer satisfaction. *Annals of Tourism Research* 38(3): 1009–1030.

Nunnally J (1978) *Psychometric Theory*, 2nd ed. New York: McGraw-Hill.

Núñez-Serrano JA, Turrion J and Velázquez FJ (2014) Are stars a good indicator of hotel quality? Asymetric information and regulatory heterogeneity in Spain. *Tourism Management* 42: 77–87.

O'Cass A and Viet Ngo L (2007) Market orientation versus innovative culture: two routes to superior brand performance. *European Journal of Marketing* 41(7/8): 868–887.

Oliver RL (1997) *Satisfaction: A Behavioural Perspective on the Consumer*. New York: McGraw-Hill.

Pantelidis IS (2010) Electronic meal experience: a content analysis of online restaurant comments. *Cornell Hospitality Quarterly* 51: 483–491.

Parasuraman A, Zeithaml V and Berry LL (1988) SERVQUAL: a multiple-item scale for measuring customer perceptions of service quality. *Journal of Retailing* 64(1): 12–40.

Prebensen NK, Woo E, Chen JS, et al. (2012) Motivation and involvement as antecedents of the perceived value of the destination experience. *Journal of Travel Research* 52(2): 253–264.

Qureshi I and Compeau D (2009) Assessing between groups differences in IS research: a comparison of covariance and component-based SEM. *MIS Quarterly* 33(1): 199–216.

Rodríguez-Díaz M and Espino-Rodríguez TF (2008) A model of strategic evaluation of a tourism destination based on internal and relational capabilities. *Journal of Travel Research* 46: 368–380.

Rodríguez Díaz M, Espino Rodríguez TF and Rodríguez Díaz R (2015) A model of market positioning based on value creation and service quality in the lodging industry: an empirical application of online customer reviews. *Tourism Economics* 21(6): 1273–1294.

Roldán JL and Sánchez-Franco M (2012) Variance-based structural equation modeling: guidelines for using partial least squares in information systems research. Research methodologies, innovations and philosophies. In: IGI Global (ed) *Software Systems Engineering and Information Systems*. Hershey: Information Science Reference, pp 193–221.

Ryu K and Han H (2010) Influence of the quality of food, service, and physical environment on customer satisfaction and behavioural intention in quick-casual restaurants: moderating role of perceived price. *Journal of Hospitality & Tourism Research* 34(3): 310–329.

Rust RT and Oliver RL (1994) Service quality: insights and managerial implications from the

frontier. In: Rust RT and Oliver RL (eds) *Service Quality: New Directions in Theory and Practice*. Thousand Oaks: Sage Publications, pp. 1–19.

Sweeney JC, Soutar GN and Johnson LW (1999) The role of perceived risk in the quality-value relationship: a study in a retail environment. *Journal of Retailing* 75(1): 77–105.

Tajzadeh-Namin A (2012) A review on value creation in tourism industry. *Management Science Letters* 2: 203–212.

Tanford S, Baloglu S and Erdem M (2012) Travel packaging on the Internet: the impact of pricing information and perceived value on consumer choice. *Journal of Travel Research* 5(1): 68–80.

Tenenhaus M, Vinzi VE, Chatelin YM, et al. (2005) PLS path modeling. *Computational Statistics & Data Analysis* 48(1): 159–205.

Varini K and Sirsi P (2012) Social media and revenue management: where should the two meet? *Journal of Technology Management for Growing Economies* 3(1): 33–46.

Vermeulen IE and Seegers D (2009) Tried and tested: the impact of online hotel reviews on consumer consideration. *Tourism Management* 30: 123–127.

Williams P and Soutar GN (2009) Value, satisfaction and behavioural intentions in an adventure tourism context. *Annals of Tourism Research* 36(3): 413–438.

Yacouel N and Fleischer A (2012) The role of cybermediaries in reputation building and price premiums in the online hotel market. *Journal of Travel Research* 51(2): 219–226.

Ye Q, Li H, Wang Z, et al. (2014) The influence of hotel price on perceived service quality and value in e-tourism: an empirical investigation based on online traveller reviews. *Journal of Hospitality & Tourism Research* 38(1): 23–39.

Zeithaml VA (1988) Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing* 52(July): 2–22.

Zhang ZQ, Ye Q, Law R, et al. (2010) The impact of e-word-of-mouth on the online popularity of restaurant: a comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management* 29: 694–700.