A Comparative Study on Hotel Consumers' Online Feedback Behaviors

Full Paper

Xianfeng Zhang
School of Information Science & Technology
Hainan Normal University
Haikou, China
xfzhangchina@gmail.com

Hongxiu Li
Turku School of Economics
University of Turku
Turku, Finland
hongli@utu.fi

Zhangxi Lin
The Rawls College of Business Administration
Texas Tech University
Lubbock, USA
zhangxi.lin@ttu.edu

Abstract

This paper is intended to analyze the interrelationship between the textual comment and the itemized rating in consumers' online feedback, and to study differentiations among diversified segments of consumers who leave the feedback. The data were collected from both an international top tourism service website which operates hotel booking services internationally and offers online hotel reputation feedback services, and its China host website. Based on the coded comments and itemized ratings from the dataset, our study shows that different customer segments demonstrate varied interrelationship behavioral patterns, and factors like satisfaction and contextual backgrounds will matter. The findings from this research may help hotels to develop varied marketing strategies for different segment of customers, and help online reputation sites to improve their services by distinguishing dissimilar behavioral patterns in different customer segments.

Keywords

Electronic word of mouth, online rating, online review, electronic commerce, tourism

Introduction

With the rapid development of the Internet and the popularity of social media, user-generated content (UGC) shoot up in the past years. Consumers produce UGC in the form of textual comment via various online feedback channels, mainly websites, to share their prior consumption experiences with others (Li and Liu 2014; Litvin et al. 2008). This kind of UGC effectively helps other consumers to make better purchasing decisions, and greatly reduces information asymmetry in e-commerce (Alkerlof 1970; Li and Liu 2014). In addition to online textual comment, majority of e-commerce websites allow consumers to rate a service or product they purchased in terms of a few given aspects. The quantitative information of the online rating (mostly including overall rating and multiple itemized rating) further eases others in decision making. As the online feedback system with the features of both textual comment and quantitative rating can effectively promote online sales in e-market, it has attracted the major attention of both researchers and practitioners.

Prior research has examined the impact of online feedbacks on sales or consumers' decisions from different perspectives, such as the value and variance of online rating, and the valence or volume of textual comment (Hu et al. 2014; Tsang and Prendergast 2009). Though the literature has acknowledged the importance of both online textual comment and online rating, little research has tried to explore the interplay of the two sets, which reflects reviewers' behavioral justifications. More specifically, even fewer discuss the relationship from the itemized components, for example, the consistency of the online textual comment and the online itemized rating in each feedback. In addition, the contextual background of those who produce both online textual comment and itemized rating has been overlooked in the prior research. This raises the need to study the relationship between online textual comment and online itemized rating in each feedback among different consumer segments, in particular, whether consumers' contextual background affect the consistency between the corresponding interplay.

In this research, we investigate the interrelationship between online textual comment and online itemized rating among different consumer segments in the research context of hotel service. We collected the empirical data from both an international top tourism service website and its China host website, which offers online hotel itemized ratings accompanied with online textual comments. In doing so, our research can contribute to the understanding of the interplay between the two sets, and to reveal the online eWOM behavior of consumers with empirical validation. In addition, this study will help to understand the role of contextual background in consumers' online feedback.

The remainder of this article is organized as follows: the research hypotheses and research model are introduced in the section two. The research methodology is presented in section three, followed by the discussion of the research results and the research findings in the fourth section. The paper then highlights the contribution of the current research for both research and practice, and points out the limitations for future research in the final section.

Research Model and Research Hypotheses

Research on online review and online rating

There has been increasing research on the importance of online textual comments on sales, especially considering the difference of online review volumes and review valence. Huang and Chen (2006) investigate the impact of the online review valence on sales, and find that negative online review has a negative influence on product evaluation and sales. Ghose and Ipeirotis (2011) investigate the impact of online review on sale from the lens of the review textual characteristics, and conclude that the subjectivity, informativeness, readability and linguistic correctness of the online textual comment influence its perceived usefulness and sales, whereas online textual comment mixed with objective as well as highly subjective sentences affect sales negatively. Another important research stream is the online itemized rating. Prior researches find that online customer's itemized rating affects sales. Ogut and Tas (2012) explore the impact of itemized rating on hotel sales, and result in the positive correlations. Ye et al (2011) adopt the variance of the rating, and also find the significant role of influencing hotel sales.

Recently, researchers attempt to explore the interplay of online textual comment and online overall rating, and try to distinguish the different roles of online textual comment and online overall rating in supporting consumers' decision. Hu et al. (2014) empirically explore the interactive role of online *review and online overall rating towards sales in the research context of online book sales, and validate that online overall rating does not affect sales directly, but indirectly via sentiments, and sentiments have significant direct influence on sales. Tsang and Prendergast (2009) investigate how the valence of textual comments and overall rating affects consumers' reaction, and test that the valence of online textual comment affects the perceived interestingness and trustworthiness of reviews and purchasing intention, and there is an interaction between textual comment and overall rating that affects the trustfulness of online textual comment.*

This study aims to explore the interrelationship between the two constructs- rating system that can directly capture consumers' perceptions through adopting numerical numbers (RATE) and the comment system which facilitates a free space for consumers to express their understandings (COMMENT). Unlike the prior researches that investigate the interactive relationship between the two constructs towards online sales, this study explore their inter-consistency among different consumer segments, and can further constribute to the marketing strategy for varied consumer groups.

Research Hypotheses

The cognitive consistency theory (Festinger 1957) suggests that people are motivated to change and act consistently with their beliefs and perceptions. Therefore, customers who bear a comprehensive understanding of the overall hotel service quality will express consistent opinions according to their perceptions of the varied aspects of the hotels, no matter by means of rating in numerical numbers with the scale of 1 to 5, or leaving text comments underlying their sentiments. From this regard, ratings and the relative comments should deliver the same message. This understanding has been tested truth in validating the consistent reputational power of binary ratings and itemized ratings in C2C marketplace (Zhang et al 2012). Therefore, the following hypothesis is proposed:

H1: RATE and COMMENT are significantly correlated through providing consistent signals.

There should still be some variations underneath the general consistent relationship between the two sets of opinions. First of all, due to the diversified contextual background and the varied prior experience, consumers may possess different capabilities to understand, perceive or evaluate their experience in the hotels. Accordingly, their post-transactional rate or review will eventually differ in the valence and variation. The valence refers to the reviewers' perception of positive or negative feelings, while variation mainly represents their evaluation of the degree adhered to the valence. When the reviewers cannot well capture the valence or variation of certain aspects during RATE or COMMENT, they may choose to leave it blank. From this aspect, it is reasonable to understand that there should be diversified customer patterns bearing dissimilar degrees of interrelationship between the two constructs. Therefore, the following hypothesis is proposed.

H2: There are dissimilar degrees of interrelationship between RATE and COMMENT to underlie the behavioral patterns of varied customer segments.

According to the prospect theory (Kahneman and Tversky, 1979), people are inclined to value more weights on the unhappy experience which may encounter potential losses, than on the pleasure events with the equivalent amount. The negative ratings have been widely verified to play comparatively more important role than the positive ratings, consequently engendering higher *performance (Ba and Pavlou 2002). In this regard, customers who are highly dissatisfied with the general service of the hotels will be likely to generate impressive perceptions, and consequently will be inclined to under-rate the respective aspects, and write text comments full of negative sentiments. Comparatively, satisfied reviewers will be less likely to carefully retrieve the memory and produce ratings or comments in details. Therefore, H2a is proposed.*

H2a: Comparing with satisfied customers, the dissatisfied have higher tendencies to touch details in both RATE and COMMENT, therefore generating closer significantly relationships.

Generally, satisfied customers are inclined to care less about the details as compared to the dissatisfied. However, when customers stay in the top star hotels which have been deemed as a luxury in the normal sense, they are likely to rate in a recognized norm when there are no extremely unhappy experience happened, as it is indicated by the social cognitive theory (Bandura 1986), which explain how people remember the sequence of events and use this information to guide their own subsequent behaviors through replicating the behavior model. In addition, due to the limited processing capacity, customers are also inclined to purposely reduce the amount of effort that required for them to rate and comment. Since the cognitive processing routes through which one rate and comment also vary (Nan 2014), when customers are annoyed by the enormous information load in either of the routes, they may choose to carelessly rate, or randomly leave some words bumped in their heads. In this way, the interrelationship between RATE and COMMENTS will show loose correlation or even confounding connections. Therefore H2b is proposed:

H2b: There are loose or inconsistent connections between RATE and COMMENT for some satisfied customers.

People from different context background will show dissimilar behavioral tendencies. The trend is found in either RATE or COMMENT system, among which Chinese customers show unique behavioral features. Zhang et al (2012) find Chinese raters cannot signal the truth in the binary rating system, but are bolder to expressing their opinions under an anonymous itemized rating system. Koh et al. (2010) confirm that Chinese reviewers normally write fewer extremely negative movie reviews. Therefore, it is proposed:

H3. Chinese customers will show behavioral difference in either the two systems - RATE or COMMENT (H3a), or their interrelationships between the two constructs (H3b), when comparing with customers from other contextual backgrounds.

Research Methodology

Data Collection

We collected empirical data on a website specialized on UGC in tourism serivces - Tripadviser. Due to the language barriers, only tripadvisor International (English) and tripadvisor China (Chinese), localized as Daodao.com, are finally selected. Top star hotels (4 or 5 star) in one popular tourism destionation, Sanya, are considered in this research for they attract the vast majority of the international customers. A two-week data crawling process started from March 23, 2014 until April 5, 2014, are conducted to collect all feedback information from top star hotels (100 in total) located in Sanya, China. Totally, 24,051 feedback records in Chinese and 21,104 tripadvisor feedback records are obtained. In cleansing the data, all feedbacks in Chinese are left for the daodao dataset, *and tripadvisor dataset is ensured to have only English records, so as to filter out the comments of other multiple languages.*

*The two samples contain the same items, including overall rating (**OverallR**), room rating (**RoomR**), service rating (**ServiceR**), value rating (**ValueR**), cleanliness rating (**CleanR**), location rating (**LocationR**), sleep quality rating (**SleepR**), and text comments.*

Content Coding of the Comments

Since this paper is still in the exploratory stage to seek for the inherent relationship between RATE and COMMENT, content analysis of a smaller sample portion is finally opted. Content analysis is a popular technique to interpret the text documents of any kind (text, oral, iconic, audio-visual) to objective data, by means of manually coding or computer-assisted procedures, so as to retrieve meaningful information, and ultimately produce the valid and trustworthy inferences (Krippendorff 1980). Due to its advantage in retrieving the rich contents, the method has been widely adopted in the IS field. A random sample containing 220-250 observations is finally derived from both datasets for content analysis.

*Systematic procedures should be followed to ensure the objectivity and reliability of the data analysis (Kolbe and Burnett 1991). Only the six analogous itemized rating components (room, service, value, cleanliness, location, sleep) are adopted as the coding category, and each being denominated as **RoomC**, **ServiceC**, **ValueC**, **CleanC**, **LocationC**, **SleepC** in the data analysis, though customers also touch multiple aspects regarding the general hotel service, like food, amenities, and etc. A five scale measurement (-2, -1, 0, 1, 2) is adopted in the coding to well reflect the varied sentiments in different categories, with 2 standing for the superb perceptions of the related aspect, while -2 describing the worst opinions. There are chances that reviewers leave no relative textual comments in certain category, a number of 0 is then given or just leave it as missing. Two independent coders who can well understand both Chinese and English language code all the comments, and the inter-rater reliability is acceptable.*

Descriptive Statistics

Since some customers do not reveal any sentiments in any of the aspects of the textual comments, which inevitably bring difficulties in exploring the interrelationship, only the observations with reviews coded in any of the six aspects are finally reserved. Consequently, 206 international observations and 240 domestic samples are left as valid in this study. As shown in the Appendix Table 1, the aspect of sleep quality attracts few to none ratings and comments. Only 6 domestic customers out of 240 put ratings in sleep quality, though the number reaches 152 among the 206 internationals, still only 12 customers express their sentiments in the textual comments. Sleep quality is therefore not considered in the following analysis. Similarly, not all five components are rated or commented by customers, indicating different people emphasize on diverse aspects which have deeply impressed them. For instance, only 50 international customers leave textual comments regarding cleanliness, whereas those who talked about service get to the number of 181. The dissimilar behavioral features can be captured by the mean and variance. Accordingly, a series of derivative variables are generated, with Mean5R and Variance5R each representing the mean and variance among the five important itemized ratings left by each customer. Missing values are not counted so as to ensure the variable can well reflect the customers' opinion. Mean5C and Variance5C *stand for the same in the case of coded textual comments. Generally speaking, the higher the mean, the positive the customers experience in their staying in the hotel, whereas the higher the variance, the more chance that dissimilar aspects of the overall hotel service impress differently on the customers. In addition, the number of the missing value among the five components is also derived as Missing5R and Missing5C.*

Data Analysis

Canonical correlation analysis is adopted for the purpose of exploring the interrelationship between the two sets of variables both underlying the customers' sentiments towards five aspects. This multivariate statistical model explores the interrelationships among two sets of dependent variables and independent variables (Grandon and Pearson 2004; Mahmood and Mann 1993). Unlike multiple regression, that predicts a single dependent variable from a set of multiple independent variables, canonical correlation simultaneously predicts multiple dependent variables from multiple independent variables (Hair et al. 1998). In the current research, the predictor variables are from the coding of the textual comments, including RoomC, ServiceC, ValueC, CleanC, and LocationC. The criterion variables are the five itemized rating, consisting of RoomR, ServiceR, ValueR, CleanR, and LocationR. SAS 9.4 is used in the canonical correlation analysis.

For robustness check, a total of eight models are conducted independently for each group of samples. Since the main research goal is to stress out the varied behavioral patterns, the eight models share the same set of predictor and criterion variables, and with only varied number of observations. Diversified rating patterns (Mean5R and Variance5R) are borrowed for the segmentation, and the mean and variance of coded comments are not used to segment customer groups, since the manually coded numerical numbers might not well reflect the customers' initial evaluations.

Higher means across the five itemized ratings represent comparatively positive evaluations, and low means accordingly stand for negative opinions. As it is commonly recognized, in a five scale measurement, 1 and 2 indicate negative opinion, whereas 4 and 5 will express positive understanding, and the neutral ratings of 3, insightful to represent the consumers' dissatisfaction (Dellarocas and Wood 2008) can be included in the segment of less mean. Therefore, model 2 (I-M2/D-M2) and model3 (I-M3/D-M3) are conducted based on the cutoff of 3 in Mean5R. The reviewers with zero variances across their itemized ratings (VarainceR5) mostly leave exactly the same or similar rating numbers in all the five components, while those of nonzero variances show discrepancies in their understanding of the five aspects, correspondingly resulting in model 4 (I-M4/D-M4) and model 5 (I-M5/D-M5).

Three extreme cases are also considered. Reviewers who possess the itemized ratings with large-mean and zero-variance (I-M6/D-M6) choose either to rate five or four in all hotel service aspects. Those satisfied customers expressing dissimilar evaluations towards the five aspects will engender the feature of large-mean nonzero-variance itemized ratings (I-M7/D-M7). Still some dissatisfactory customers are willing to share their unhappy experience, yielding a group of sample with small-mean and nonzero-variance (I-M8/D-M8). The observations with small mean and zero variance are less than 5 in both samples, that case is therefore not considered in this study.

In addition, paired and independent t-tests using SPSS 19.0 are also conducted to test the difference between the two groups of samples.

Results and Discussions

The Analyses of the General Samples

Canonical correlation analysis requires variables within one set to be correlated yet not so close to have multicollinearity issue. Following Thorndike (1978), we tested for correlation and multicollinearity before the canonical analysis. The variables of each group are found to be correlated but no multicollinearity problem is found in the preliminary test.

As suggested by Hair et al. (1998), when interpreting the output of the canonical correlation analysis, a first step is to test whether there is any significant relationship between the two sets of variables– predictor and criterion. It is normally judged by four multivariate tests of significance (Wilks' Lambda, Hotelling-Lawley Trace, Pillai's Trace, Roy's Greatest Root), by which Wilks' Lambda that also equals to the Pr>F value in the first model, is the most important (as reported in the 3rd row of table 2-3). It is found that, taken collectively, the canonical functions in six out of the eight models (except M5-6) of both customer groups, are statistically significant at the 0.01 level, showing RATE and COMMENT variables are significantly correlated. H1 is supported. The result is consistent with the cognitive consistency theory (Festinger 1957), implying that customers will normally act consistently with perceptions regarding the general hotel service. Therefore, the information derived from the customers' textual comments should be similar to what they have left in the rating system, regardless of which measurement they are using.

Second, the significant test of the canonical correlation needs to be interpreted. This tells which canonical functions can fit well, and whether the canonical variables significantly correlate. According to Green, et al (1966), the maximum number of canonical functions is the number of variables in the smaller set. As in this study, there are five canonical functions generated. Table 2 and Table 3 report only the first two canonical functions for short of space. It is obvious that in all the 10 significant models, indexes of the canonical correlation and squared canonical correlation, are larger in the first canonical function as it is in the second, indicating a much better fitness of the first function. As well, the first function also takes lower or at least similar significance (Pr>F). It has also been recommended that redundancy analysis is better to confirm which function fits better (Hair 1998; Grandon and Pearson 2004). Redundancy analysis tells us how much of the variance in one set of variables can be explained by another set of canonical variables. The redundancy index, analogous to the R^2 statistic of the multiple regression is derived by multiplying the two components - shared variance of the covariate (the variance of one set of variables explained by the other set of variables) and the squared canonical correlation (Haier et al. 1998). Take I-M1 in Table 2 for instance: 36.2% of the variance of the dependent and independent covariates is explained by the opposite variables in the first canonical function, far larger than that of 2.9% in the second function. The same pattern holds for other models. Therefore, only the first canonical function is considered for interpretation in all the models.

The canonical structures show how the internal variables in either the predictor or criterion construct correlate with the opposite covariate. Cross-loadings are reported in order to increase the stability and external validity of the findings (Hair et al. 1998). Using a cutoff correlation of 0.3, recommended by Lambert and Durand (1975) as an appropriate value for structure coefficients, all the significant variables are marked out in Table 2 and Table 3. It is interesting, yet not surprising to find that customers of dissimilar ratings, which indicate the diversified behavioral patterns, *engender totally different canonical structures. As in both Chinese context or in the international sphere, customers rate lower or rate variedly in different components induce higher canonical structures than those rating positively in a general way. Therefore H2 is supported, indicating that there exist diversified feedback behavioral patterns.*

No relative relationship is found between RATE and COMMENT for the customers who contain zero variance among their itemized ratings. The findings make sense when counting the fact that those customers who quickly put all the same numbers in the five itemized rating components may write a few words, either complaining or praising certain aspects. A set of same value variables can hardly have positive correlation with the other dataset that contains varied coding of contextual comments. In order to well understand the phenomena, extreme examples are studied consequently.

Analyses of the Extreme Samples

Three extreme cases (M6, M7 and M8) are also studied for both domestic and international samples. The customers with small-mean and nonzero-variance rating pattern (I-M8 and D-M8) show the strongest relationship either between the two sets of variables, or within the sets. Its explanation power, denominated by the redundancy index ranks almost the highest among the eight models, although the redundancy index explained by the international reviewer observations with nonzero variances take a marginal advantage. The canonical structures of the two models also highlight the strong positive connections between itemized rating and textual comment. However, the other two groups (I-M7 and D-M7) with nonzero-variance and large-mean pattern do not display strong canonical structures. International customers express significant correlation in aspect of service, whereas domestic customers show a marginal significant relationship in the aspect of value.

This finding support H2a, evidencing that dissatisfied customers tend to stress more efforts during rating and commenting, and go into the details in both measures, while the satisfied customers will address comparatively less attention. It is comply with the Prospect theory (Kahneman and Tversky, 1979) which indicates that customers appear to give higher value to the unhappy experience of loss than to the pleasant memory in the hotel. Not only do they rate very low, they also write long comments describing the annoying process, and pinpoints the evaluation towards some or all of the general hotel service.

As predicted, no canonical correlations between RATE and COMMENT are found in both customer segments leaving similar positive ratings with zero variances (large-mean zero-variance pattern). It indicates that ratings are not correlated with the signals derived from text comments for those customers. Surprisingly, the canonical structures in both samples express certain negative relationships, for instance, respectively four (three) out of the five factors among the international reviewers (domestic reviewers) are negatively related, which underlies the confounding signals between itemized rating and textual comment. Therefore H2b is supported. As well, cognitive load theory and social cognitive theory should help explain the phenomena, as it is pointed out earlier.

Compared Analyses between Samples under Different Context

The customer from Chinese contextual background has been tested to show unique behavioral features than those from other background. Therefore the study roughly deems the international review dataset as one unique group bearing the international viewpoint, though the international customers may further be classified into dissimilar cultural background.

Generally, there is no significant difference in the means of the five component itemized ratings[1] between local Chinese reviewers and international reviewers. Difference (though marginally at the sig level of 0.091) is found in the means of the textual comment coding between the two samples with a 0.486 margins. However, international reviewers possess higher variance in either the itemized rating (0.444) or the textual comments (0.291).

The independent t-test is then conducted through adopting the variables that can well distinguish the different reviewers' behavior. Each reviewer is accounted as an independent sample. As it is showed in Table 4, the prior findings are reinforced. There are no significant differences in the means of the itemized rating or textual comment coding between the two samples, whereas the international customers tend to show variances in their expression of the opinions. This is conformed to the prior research (Koh et al. 2010; Zhang et al 2012), implying that deeply rooted in the Confucian philosophy, Chinese customers tend to treat others as what they are preferred. This tendency is also reflected through the significant difference in the number of missing itemized rating. Therefore, the H3a are supported.

Interestingly, it is also found that domestic customers and international customers care differently about the components of hotel service. Among those reviewers with lower itemized ratings, international customers stress details about service, value, cleanliness and room, whereas domestic customers mostly only care about room and service. For the international reviewers with pleasant experience, the aspect of service attracts significant high correlation among the two sets of measurement, and domestic customers pinpoint in the value aspect.

Concluding Remarks

This paper explores the interplay between the two commonly adopted feedback information - the most important online reputation signals – itemized rating which measures the consumers' perception of the sellers' general performance in multiple aspects by adopting numerical numbers, and textual comment, which facilitates the delivery of the relative understandings in a free format. It is found that the two sets of constructs - RATE and COMMENT, are displaying consistent information. Further differentiated by means and variances of the itemized ratings, customers are segmented into diversified groups to discover their review behavioral patterns, and the interconnection between itemized rating and contextual comment is then investigated for varied customer segments. It is concluded that dissatisfied customers tend to have closer interplay between RATE and COMMENT, and the interrelationship appear loose or inconsistent for satisfied customers. In addition, from the cross-contextual perspective, the research comparatively distinguishes the Chinese customers' WOM behaviors and their interrelationship between RATE and COMMENT, from those of customers from other contextual background. The conclusion reinforce the findings of prior literature, indicating that Chinese customers show behavioral difference in either the two systems - RATE or COMMENT, or their interrelationships between the two constructs.

Through connecting itemized rating and textual comment by the adoption of canonical correlation analysis, the research is intended to enrich the research literature, by validating the general

[1] The means of the five component ratings (RoomR, ServiceR,ValueR, CleanR, LocationR) as depicted in the Appendix Table 1 are adopted as the observations for the paired t-test. The same rule applies to the other paired t-test variables.

significantly consistent interrelationship between the two constructs, though the varied groups show diversified performance. It can help understanding the consistent reputation signals and the consumers' eWOM behavior, and also elucidate the role of contextual backgrounds. In addition, the research may help the enhancement of rating items in the hotel eWOM platform to better reflect the service qualities of the hotels, and to capture the customers' perceptions towards different aspects.

Practically, evidence shows that different customer segments demonstrate varied behavioral characteristics. Hotels could improve their services by specifically analyzing those customers with low itemized rating and high variance. When dealing with customers of different contexts, it is better to learn the differences from the satisfied customers bearing high itemized rating and high variance. Specifically take the hotel industries of a hot leisure city in China for instance, great care need to be put for international customers in the aspect of hotel services such as fluent English, friendliness, and professionalism, while local Chinese customers demand good values like lower price. Therefore, hotels need to develop varied marketing strategies for different customer segments. The last but the most important, there are some imperfects in the current reputation systems. Many of them could not distinguish dissimilar behavioral patterns in different customer segments, so as to less effectively encourage their consistent generation of eWOM in different systems.

Due to limited available time for data processing, we only used 446 observations randomly sampled from a large dataset with manual coding. The small size of the dataset limited the subgroup discussion, especially in the case of the group with zero rating variances. Accordingly, the external validity was affected. Fortunately, major interesting findings are discovered in this explorative study, which will provide confidence in the future analysis using the dataset in larger size.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Grant No.71362027), and MOE Humanities and Social Sciences Project of China (No.13YJC630228).

REFERENCES

Akerlof, G. 1970. "The market for lemons: Quality under uncertainty and the market mechanism," Quarterly Journal of Economics (84), pp. 488–500.

Bandura, A. 1986. Social Foundations of Thought and Action : A Social Cognitive Theory, Englewood Cliffs, N.J.: Prentice-Hall.

Festinger, L. 1957. A Theory of Cognitive Dissonance. Stanford, CA: Stanford University Press.

Goose, A. and Ipeirotis, P.G. 2011. "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics," IEEE Transactions on Knowledge and Data Engineering (23:10), pp. 1498-1512.

Grandon, E. E., and Pearson, J. M. 2004. "Electronic commerce adoption: an empirical study of small and medium US businesses," Information & Management (42), pp. 197–216.

Green, P. E., Halbert, M. H., and Robinson, P. J. 1966. "Canonical analysis: an exposition and illustrative application," Journal and Marketing Research (3), pp. 32–39.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. 1998. Multivariate Data Analysis. New Jersey: Prentice-Hall.

Hu, N., Koh, N.S. and Reddy, S.K. 2014. "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," Decision Support Systems (57), pp. 42-53.

Huang, J.H. and Chen, Y.F. 2006. "Herding in online product choice," Psychology and Marketing (23:5), pp. 413-428.

Kahneman, D., and Tversky, A. 1979. "Prospect theory: an analysis of decision under risk," Econometrica (47:2), pp. 263–292.

Kolbe, R. H., and Burnett, M. S. 1991. "Content analysis research: An examination of applications with directives for improving research reliability and objectivity," Journal of Consumer Research (18:3), pp. 243–250.

Krippendorff, K. 1980. Content Analysis: An Introduction to Its Methodology. Sage Publications, CA: Beverly Hills.

Lambert, Z.V. and Durand, R.M. 1975. "Some precautions in using canonical analysis,"Journal of Marketing Research (12: November), pp. 468-475.

Li, H. & Liu, Y. 2014. "Understanding post-adoption behaviors of e-service users in the context of online travel services," Information & Management (51:8), pp. 1043-1052.

Litvin, S.W., Goldsmith, R.E., and Pan, B., 2008. "Electronic word-of-mouth in the hospitality and tourism management," Tourism Management (29:3), pp. 458-468.

Mahmood, M. A., and Mann, G. J. 1993. "Measuring the organizational impact of information technology investment: an exploratory study," Journal of Management Information Systems (10:1), pp. 97-122.

Ogut, H. and Tas, B. 2012. "The influence of internet customer reviews on the online sales and prices in hotel industry," The Service Industries Journal (32:2), pp. 197-214.

Thorndike, R. M. 1978. Correlational Procedures for Research. New York: John Wiley & Sons Inc.

Tsang, A., and Prendergast, G. 2009. "Is a "star" worth a thousand words? The interplay between product-textual comments and rating valences," European Journal of Marketing (43:11/12), pp.1269-1280.

Ye, Q., Law, R., Gu, B., and Chen, W. 2011. "The influence of user-generated content on customer behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings," Computers in Human Behavior, (27), pp. 634-639.

Zhang, X., Luo, J. and Li, Q. 2012. "Do different reputation systems provide consistent signals of seller quality: a canonical correlation investigation of Chinese C2C marketplaces," Electron Markets (22), pp. 155-168.

Appendices:

Table 1. Descriptive Statistics

International Reviewers	MIN.	MAX.	MEAN.	S.D.	Domestic Reviewers	MIN.	MAX.	MEAN.	S.D.
Ratings					Ratings				
RoomR (N=187)	1	5	4.28	.971	RoomR (N=240)	1	5	4.11	.697
ServiceR (N=206)	1	5	3.83	1.338	ServiceR (N=240)	1	5	4.15	.704
ValueR (N=206)	1	5	3.74	1.192	ValueR (N=240)	1	5	3.95	.651
CleanR (N=206)	1	5	4.27	1.019	CleanR (N=240)	1	5	4.19	.609
LocationR (N=184)	1	5	4.08	.989	LocationR (N=240)	2	5	4.20	.626
SleepR (N=152)	1	5	4.24	.983	SleepR (N=6)	4	5	4.5	.548
Missing5R (N=206)	0	2	.180	.552	Missing5R (N=240)	0	0	.00	.000
Mean5R (N=206)	1	5	4.030	.899	Mean5R (N=240)	1.40	5	4.119	.529
Variance5R (N=206)	0	3.20	.596	.684	Variance5R (N=240)	0	2.80	.202	.291
Review Sentiments					Review Sentiments				
RoomC (N=144)	-2	2	.74	1.177	RoomC (N=175)	-2	2	.80	1.114
ServiceC (N=181)	-2	2	.24	1.622	ServiceC (N=195)	-2	2	.98	1.074
ValueC (N=82)	-2	2	-.10	1.375	ValueC (N=75)	-2	2	-.19	1.159
CleanC (N=50)	-2	2	.04	1.414	CleanC (N=82)	-2	2	1.13	.926
LocationC (N=80)	-2	2	.13	1.267	LocationC (N=185)	-2	4	.76	1.127
SleepC (N=12)	-2	1	-1.00	1.044	SleepC (N=9)	-2	2	0	1.5
Missing5C (N=206)	0	4	2.136	1.265	Missing5C (N=240)	0	4	2.10	.932
Mean5C (N=206)	-2	2	.352	1.185	Mean5C (N=240)	-2	2	.755	.843
Variance5C (N=171)	0	8	1.216	1.385	Variance5C (N=224)	0	4.5	1.048	1.134

Table 4. Reviewer's Behavior between International and Domestic Reviewers

	Overall rating D/I	Missing 5R D/I	Missing 5C D/I	Mean 5R D/I	Variance 5R D/I	Mean 5C D/I	Variance 5C D/I
N	240/205	240/205	240/205	240/205	240/205	240/205	224/171
Mean	4.19/3.94	0/0.18	2.10/2.15	4.12/4.03	0.20/0/60	0.76/0.35	1.05/1.22
S.D.	0.63/1.13	0/0.533	0.93/1.26	0.53/0.90	0.29/0.69	0.84/1.18	1.34/1.39
F	8.64	25.6	0.198	1.702	65.07	1.75	17.48
Sig	**0.003**	**0.000**	0.657	0.193	**0.000**	0.186	**0.000**

Table 2. Canonical Analysis Result of International Reviewers

Tests or variables	*I-M1 Overall (N=206)*	*I-M2 Large mean (N=133)*	*I-M3 Small mean (N=73)*	*I-M4 Nonzero variance (N=169)*	*I-M5 0-variance (N=37)*	*I-M6 0-variance large mean (N=34)*	*I-M7 large mean Nono-variance (N=99)*	*I-M8 small mean Nono-variance (N=70)*
Canonical Corr.	0.820/0.5254	0.662/0.399	0.777/0.620	0.823/0.566	0.631	0.598	0.723/0.456	0.790/0.683
Canonical R2	0.672/0.276	0.438/0.159	0.603/0.385	0.677/0.320	0.398	0.358	0.522/0.208	0.624/0.466
Pr>F	<0.0001/<0.0001	<0.0001/0.002	<0.0001/<0.0001	<0.0001/<0.0001	0.125	0.252	<0.0001/<0.000	<0.0001/<0.000
Shared Var. (C)	0.226/0.043	0.084/0.035	0.185/0.064	0.220/0.051	0.106	0.069	0.103/0.051	0.201/0.064
R. index(C)	0.152/0.012	0.037/0.006	0.112/0.025	0.149/0.016	0.042	0.025	0.054/0.005	0.125/0.030
Shared Var. (R)	0.313/0.026	0.083/0.018	0.201/0.040	0.293/0.037	0.275	0.148	0.109/0.027	0.246/0.084
R. index (R)	0.210/0.007	0.036/0.003	0.121/0.015	0.198/0.012	0.109	0.053	0.057/0.007	0.154/0.039
RoomR	0.437** (0.191)	-0.118 (0.014)	0.611** (0.373)	0.443** (0.196)	0.245	0.317	-0.129 (0.017)	0.637** (0.405)
ServiceR	0.793** (0.629)	0.598** (0.358)	0.551** (0.303)	0.793** (0.629)	0.591	0.394	0.638** (0.407)	0.564** (0.319)
ValueR	0.356** (0.127)	-0.102 (0.010)	0.362** (0.131)	0.322** (0.104)	0.268	-0.008	-0.157 (0.025)	0.438** (0.192)
CleanR	0.427** (0.182)	0.172 (0.029)	0.343** (0.118)	0.413** (0.171)	0.225	0.032	0.208 (0.043)	0.278* (0.078)
LocationR	0.046 (0.002)	-0.079 (0.006)	0.012 (0.000)	0.027 (0.001)	-0.020	0.295	-0.101 (0.010)	0.097 (0.009)
RoomC	0.309** (0.095)	-0.118 (0.014)	0.287* (0.082)	0.301** (0.091)	0.373	-0.157	-0.187 (0.035)	0.569* (0.324)
ServiceC	0.805** (0.648)	0.631** (0.398)	0.604** (0.365)	0.802** (0.644)	0.593	-0.488	0.693** (0.481)	0.612** (0.375)
ValueC	0.659** (0.434)	-0.003 (0.000)	0.579** (0.335)	0.618** (0.381)	0.593	-0.488	-0.096 (0.009)	0.597** (0.356)
CleanC	0.608** (0.370)	0.048 (0.002)	0.453** (0.205)	0.587** (0.345)	0.593	-0.488	0.040 (0.002)	0.419** (0.175)
LocationC	0.138 (0.019)	0.046 (0.002)	-0.122 (0.015)	0.056 (0.003)	0.427	-0.021	0.125 (0.016)	-0.416 (0.002)

** means a cross-loading above the cut-off of 0.3(Lambert and Durand 1975), the numbers in the parenthesis are the SMC. * means larger than 0.20.

Note: The first indexes before the dash comes from the first canonical functions, and the latter index after the dash is for the second canonical functions. When there is no significant relationship between the two sets of variables, only indexes from the first function are reported.

Table 3. Canonical Analysis Result of Domestic Reviewers

Tests or variables	D-M1 Overall (N=240)	D-M2 Large mean (N=178)	D-M3 Small mean (N=62)	D-M4 Nonzero variance (N=132)	D-M5 0-variance (N=108)	D-M6 0-variance large mean (N=105)	D-M7 large mean Nono-variance (N=73)	D-M8 small mean Nono-variance (N=59)
Canonical Corr.	0.532/0.298	0.357/0.318	0.707/0.326	0.669/0.394	0.190	0.253	0.534/0.465	0.713/0.367
Canonical R2	5.06/2.66	0.127/0.101	0.499/0.106	0.447/0.195	0.036	0.064	0.285/0.216	0.508/0.135
Pr>F	<0.0001/0.0004	0.001/0.001	0.003/0.850	<0.0001/0.001	0.581	0.249	0.002/0.037	0.002/0.689
Shared Var. (C)	0.071/0.016	0.023/0.019	0.134/0.018	0.122/0.028	0.008	0.013	0.035/0.061	0.139/0.024
R. index(C)	0.020/0.001	0.003/0.002	0.067/0.002	0.055/0.005	0.0003	0.001	0.010/0.013	0.071/0.003
Shared Var. (R)	0.120/0.021	0.021/0.015	0.136/0.020	0.185/0.033	0.036	0.064	0.068/0.036	0.135/0.026
R. index (R)	0.033/0.002	0.003/0.002	0.068/0.002	0.083/0.006	0.001	0.004	0.019/0.008	0.069/0.004
RoomR	0.402** (0.161)	-0.1000 (0.0010)	0.464** (0.216)	0.556** (0.309)	0.015	-0.017	-0.075 (0.0056)	0.474** (0.225)
ServiceR	0.417** (0.174)	0.124 (0.0156)	0.530** (0.281)	0.492** (0.242)	0.099	-0.031	-0.160 (0.026)	0.534** (0.285)
ValueR	0.054 (0.003)	0.243* (0.086)	0.101 (0.010)	0.042 (0.002)	-0.022	0.016	-0.222* (0.049)	0.106 (0.011)
CleanR	0.121 (0.015)	-0.046 (0.002)	0.377** (0.142)	0.232* (0.054)	-0.085	-0.203	0.091 (0.008)	0.380** (0.145)
LocationR	-0.040 (0.002)	-0.052 (0.003)	-0.146 (0.021)	-0.029 (0.001)	0.146	0.148	0.296 (0.087)	-0.170 (0.029)
RoomC	0.481** (0.235)	-0.049 (0.002)	0.625** (0.391)	0.640** (0.410)	0.190	0.253	-0.190 (0.036)	0.623** (0.389)
ServiceC	0.373** (0.139)	-0.084 (0.007)	0.417** (0.174)	0.479** (0.230)	0.190	0.253	0.014 (0.000)	0.420** (0.168)
ValueC	0.290* (0.084)	0.244* (0.060)	0.128 (0.016)	0.326** (0.106)	0.190	0.253	-0.264* (0.070)	0.113 (0.013)
CleanC	0.273* (0.075)	-0.089 (0.008)	0.271* (0.074)	0.393** (0.154)	0.190	0.253	0.131 (0.216)	0.253* (0.064)
LocationC	0.124 (0.015)	-0.159 (0.025)	-0.158 (0.025)	0.159 (0.025)	0.190	0.253	0.465 (0.025)	-0.198 (0.039)