

DO DOMESTIC AND INTERNATIONAL CUSTOMERS BEHAVE ALIKE IN ONLINE HOTEL BOOKING?

Asunur Cezar

Department of Business Administration
TOBB University of Economics and Technology
Söğütözü Cad. No:43 Ankara, TURKEY, 06560
acezar@etu.edu.tr

Hulisi Ögüt

Department of Business Administration
TOBB University of Economics and Technology
Söğütözü Cad. No:43 Ankara, TURKEY, 06560
hogut@etu.edu.tr

ABSTRACT

The aim of this research is to explore online hotel booking behaviors of domestic and international customers. We examine the impact of review rating (location rating and service rating), review volume and hotel characteristics on the proportion of online domestic bookings. Using data from London and Paris hotels, we evaluate a variety of models to explain fractional dependent variables. Our results show that price, review volume and location rating have a significant and negative impact while service and star rating have a significant and positive impact on the proportion of domestic bookings. However, the proportion of international bookings does not increase if the hotel belongs to a chain group. Moreover, domestic customers are more sensitive to changes in price and service rating whereas international customers are more responsive to changes in location rating. The results suggest that that domestic and international customers' online hotel booking behaviors show significant differences.

Keywords: Online customer reviews and ratings; Domestic and international customer; Online hotel booking; Information asymmetry

1. Introduction

With the help of recent developments in information and communication technology, more and more people use the Internet for a variety of reasons. People also increasingly prefer online sites for buying goods and services and internet becomes main sales and marketing channel for different industries. Travel industry is one of the earliest and established industries to adopt online channels for this purpose and the growth of online travel bookings is much faster than the total market [PhoCusWright 2011]. With the help of intermediary websites such as booking.com and expedia.com, online hotel booking has also become widely popular in hospitality industry and the share of the revenues generated through online booking has been constantly increasing [Marcussen 2008]. A report by Google and OTX finds that the Internet provides platform for prospective customers to learn experience of other travelers in addition to be the leading tool and the source for researching and booking travel [Google/IPSOS OTX 2011]. The study indicates that 45% make personal travel plans and 54% make business travel plans based on the online reviews and more and more people are sharing their own experiences in the Internet in order to guide prospective customers.

Hotel rooms are an example of experience goods whose quality can be evaluated only after purchase. For such goods, consumers usually seek information from a variety of sources to diminish the uncertainty of product quality before making purchase decisions. Although Internet provides detailed information about hotels and their environments, significant information asymmetry exists between domestic and international visitors. Specifically, international customers face more information asymmetry than domestic customers. Compared to international visitors, domestic tourists probably are more informed about the city, surroundings and the transportation system. Furthermore, as domestic visitors do not have language barrier, they can easily speak to local people in case they have problems. Thus, we can argue that domestic visitors are less constrained in their hotel choices and they can choose among a wider set of hotels. Taking advantage of being more informed about city and possibly making more frequent visits, they also may face a lower quality uncertainty. Due to the possible differences between domestic and

international visitors, domestic and international demand may also show different characteristics. However, despite the possible differences, domestic visitors are essentially similar to international visitors in terms of their expectation from hotel room services. Thus, it is important to understand the factors affecting the hotel reservation choices of domestic and international visitors so that hotel managers can better address the different needs and expectations of domestic and international customers. For this reason, we explore the similarities and differences between domestic and international visitors' online hotel booking behaviors. To analyze the behavior of different types of customers in online environment, we used the following variables available on the Internet: quality metrics such as service rating and star rating and hotel features such as price, location rating, size and whether hotel operates under chain name. Since this analysis is especially important for the touristic destinations with a large proportion of international visitors, we choose two of the most visited touristic cities in the world: London and Paris. In this way, we check the robustness of our empirical findings and determine resemblances and differences between London and Paris hotels.

Our analysis considers various models developed by prior research: a linear regression model, a quasi-likelihood model with the Bernoulli log-likelihood function and a parametric regression model with the beta distribution. Then, we check whether the assumptions of models are validated and assess the performance of models using the data retrieved from one of the largest online hotel booking sites. Through these empirical analyses, we identify how customer review rating (average location rating and average service rating) and volume, star rating and other hotel attributes such as average room price, location and chain name impact the percentage of online domestic bookings. We find that while the linear regression model's assumptions are violated, our data sets cannot reject the assumptions of other two models. In each data set, we observe that price, review volume and location rating have a significant and negative impact while service and star rating have a significant and positive impact on the proportion of domestic bookings. However, we find that the proportion of international customers does not increase if the hotel is a part of a hotel chain.

We also estimate the number of hotel bookings of domestic and international customers. We find that price has a significant and negative impact while location and service rating have a significant and positive impact on the number of domestic and international bookings. However, the impact of star rating on the number of domestic and international booking is not significant. If a hotel is part of a hotel chain, the international number of bookings decreases while the domestic number of bookings is not affected. Furthermore, domestic customers are more sensitive to changes in price and service rating while international customers are more responsive to changes in location rating.

Our study has many practical implications for hotel management. First, we measured price elasticity of international and domestic customers and showed that they differ. This result suggests that different price list should be offered to international and domestic customers. Second, we found that location is more important than price and service quality for international customers. Thus, if hotel owners especially target international customers, their hotels should be in convenient locations. Furthermore, they may charge higher prices to international customers since they are less sensitive to prices than domestic customers. This result may also imply that hotels should offer high quality services and lower prices if they are not located in convenient locations. Furthermore, using our empirical model, hotel managers can forecast their hotels' domestic and international sales mix, plan for seasonal changes in domestic and international demand and make decisions to better manage their fixed capacities. We also provide detailed analysis of domestic and international customers' responses to different factors.

The rest of the paper is organized as follows. After summarizing the relevant literature in the next section, we introduce our hypotheses and their background. Then, our data set and empirical models are explained. In section 5, we present results from the empirical models. We conclude with the discussion of implications and limitations of our work.

2. Literature Review

The drivers of online sales have been extensively studied recently. For example, Forman et al. [2008] and Chevalier and Mayzlin [2006] investigate the impact of consumer reviews and ratings on online sale of books. Ba and Pavlou [2002] find that customer ratings alleviate uncertainties between sellers and buyers in the ebay auctions and high seller ratings from previous buyers generate price premium for mitigating online transaction risk.

Recently, researchers investigate customers' hotel booking behaviors in online channels. Cantalops and Salvi [2014] classified research as review generating factors and impacts of online reviews. Our paper is more related to the second line of research that investigated the impacts of online reviews on the behavior of consumers and hotel managers. Among these studies, Gu and Ye [2014] showed that while responses to customer reviews are influential for low satisfaction customers, its effect on other customers is limited. Furthermore, they found that there is a positive association between customers' future satisfaction and whether they receive responses to their complaints. Xie et al. [2011] showed that the presence of reviewer information positively influence reviewer's credibility.

However customer's intention for online hotel booking is negatively affected from ambivalent online hotel reviews. Mauri and Minazzi [2013] found that online reviews increase both consumer's purchasing intention and expectation from hotel. However, hotel managers' responses to customer reviews negatively influence consumer's purchasing intention. Vermeulen and Seeger [2009] explored the effect of online reviews on customers' hotel selection through an experimental study. They showed that consumers' awareness for hotels is affected by both negative and positive reviews. However, only positive reviews lead to improvement in customers' attitudes toward hotels. This effect is less significant for more popular hotels. Ögüt and Taş [2012] explored the effect of star rating and customer ratings on online hotel booking and prices. They showed that higher customer ratings lead to higher online bookings. However, higher star rating is not necessarily associated with higher online bookings. Furthermore, they find that higher customer ratings generate price premium for hotels. Ghose et al. [2012] shows how social media and user generated content can be used for generating a new ranking system in product search engines. The recommendation of this hotel ranking system provides better value for the customer.

Our contributions to the literature are the followings. To the best of our knowledge, our paper is the first study to analyze how review and hotel characteristics impact domestic and international online bookings. Rather than investigating the impact of review rating as an aggregate figure, we divide the review rating into service rating and location rating. Specifically, the impact of location rating on these variables has not been investigated before. Moreover, we used real hotel prices and online booking data obtained from one of the popular online booking sites. We choose the two popular tourist destinations in order to strengthen the validity of our empirical findings.

3. Theoretical Background and Hypotheses

Hotel rooms are an example of experience goods whose quality can be assessed after purchase. For such goods and services, buyers usually search for additional information prior to purchasing decisions from a variety of sources to mitigate the uncertainty of product quality. Even though online platforms supply comprehensive information about hotels and their locations, significant information asymmetry still exists between potential customers and hotel owners. There are mainly two types of uncertainties: i) the buyer does not have any information about the seller prior to purchase (seller uncertainty), ii) the quality of the hotel can only be assessed only after customer's stay at the hotel (product uncertainty) [Ba and Pavlou 2002]. Akerlof [1970] pointed out that the informational asymmetry resulting from these uncertainties can lead to market failure. With the help of various signals such as price, star and customer rating, sellers are able to show quality of their hotels and lessen the negative effects of information asymmetry [Spence 1973].

There is also information asymmetry between domestic and international customers and the degree of information asymmetry between these two types of customers may impact their hotel choices differently. Specifically, international customers may face more information asymmetry than domestic customers as they have informational disadvantage over domestic customers. Consequently, these two customer types can give different responses to the same signal as they face different degrees of uncertainty. In this study, we explore the impact of service and location rating, star rating, review volume and hotel characteristics such as, price, chain group membership and hotel size on the proportion of domestic sales. These features may influence domestic and international sales differently. In general, domestic customers are more knowledgeable about local hotels, the city, its surroundings, transportation system, culture and society. Therefore, they may consider a larger set of hotels. However, more constrained in their choice of hotels, international customers are more likely to prefer hotels located in more convenient locations.

Especially in capacity-limited service industries such as airline industry, hotels, and rental car industry, firms implement revenue management strategies as they can segment their customers according to some customer characteristics and offer different prices to different segments. Basically, firms identify more(less) price-sensitive customers and charge them lower (higher) prices. Because of mentioned informational advantage over international customers, as domestic customers face lower information asymmetries, they can consider a larger set of hotels and prefer to pay less for hotels with comparable quality. Furthermore, earlier research on the shopping habit of domestic and international customers found that domestic customers are more dissatisfied with higher price [Yüksel 2004]. One possible reason for this is that domestic customers have more information to assess whether prices of hotel are relatively high [Yüksel 2004; Ariffin et al. 2012]. Therefore, we claim that domestic customers are more price sensitive than international visitors and we have the following hypothesis:

H1: *All else equal, higher hotel price is associated with less domestic online bookings.*

Star rating provides guidelines for prospective customers for their assessment of a hotel with a certain star. Thus, star rating is traditionally regarded as one of the most significant feature in the preference of hotel choice [Callan 1998]. Like star rating, online customer ratings can provide valuable information to prospective customers as they reflect the experience of past customers. Earlier research shows that higher online customer ratings lead to

higher online sales of goods and services [Chevalier and Mayzlin 2006]. Specifically, Ögüt and Taş [2012] showed that a 1% increase in customer rating increases online hotel bookings by up to 2.68%. Industry reports confirm these results by stating that most of the people read hotel's online reviews before their stay at that hotel and customers' hotel selection decisions are significantly influenced by user generated content [comScore and the Kelsey Group, 2007].

Although customer and star ratings give information about how good the facilities and the provision of hotel services is, unlike star rating, customer ratings also provide information about the location of a hotel. Therefore, we analyze the impact of customer rating in two main dimensions. The first dimension measures the service performance of hotels like star rating and we call it service rating. This metric summarizes the performance of hotels in hotel staff, service/facilities, cleanliness, comfort and value for money dimensions as it is shown in Figure 2. The second dimension denotes the average customer rating for hotel's location and we call it location rating. These ratings convey the same information to both domestic and international visitors and all else equal, both domestic and international customers have preference for hotels with higher service and location ratings (star rating) over hotels with lower service and location ratings (star rating). However, reviews and ratings are the expression of the subjective opinions of consumers and different consumers may rate the credibility of review and ratings differently [Pathak et al. 2010]. Hence, the customer rating may have a differential impact on domestic and international customers. Moreover, customers may trade off location and service quality for a given budget. In particular, compared to international customers, domestic customers are more informed about the city's transportation system and hotels' locations. Furthermore, domestic customers can easily communicate with local people in case they need help since they do not have cultural or language barriers. Thus, domestic customers can consider hotels in a wider area without compromising quality, while international customers are more likely to consider hotels located in more popular or more convenient locations. Consequently, domestic customers may prefer to stay at a hotel with high service (star) rating which is not located at the very heart of the city (or near to major tourists attractions) than to stay at a hotel with low service (star) rating at the city center. Therefore, we have the following hypotheses:

H2: *All else equal, higher star rating is associated with more domestic online bookings.*

H3.a: *All else equal, higher location rating is associated with less domestic online bookings.*

H3.b: *All else equal, higher service rating is associated with more domestic online bookings.*

Earlier research found that there is a positive effect of review volume on online sales [Duan et al. 2008; Zhang et al. 2010]. Higher number of reviews for a hotel shows the popularity of the hotel and creates a confidence for prospective customers as many others also make reservation in the hotel [Park et al. 2007]. Furthermore, review volume helps to reduce the concerns regarding the subjectivity of customer reviews since the review rating of hotels with higher review volume on average is less likely to convey biased information due to the law of large numbers. As international customers face higher quality uncertainty, they may pay more attention to the number of customer reviews. For this reason, we conjecture that:

H4: *All else equal, higher number of customer reviews is associated with less domestic online bookings.*

We also investigate whether hotels with chain name affect the behavior of domestic and international customers differently. Although these hotels have some differences in quality as they operate in different places and countries, they apply the same procedures and standards to meet certain quality criteria. It has been reported that branded hotels are preferred over independent hotels by most of business and leisure travelers [Yesawich 1996]. The possible explanation for this finding is that chain name helps prospective customers to mitigate the risk associated with quality uncertainty [Bharadwaj et al. 1993]. Thus, brand name enables chain hotels to differentiate themselves from other hotels [O'Neill and Xiao 2006]. International customers are more likely to prefer chain hotels as they have higher information asymmetry compared to domestic customers. Thus, we conjecture that

H5: *All else equal, hotels belonging to chain group have lower proportion of domestic online booking.*

4. Methodology

4.1. Data Overview

Source of our data is www.booking.com. Booking.com is one of the leading online hotel booking platforms. Our dataset contains online hotel reservations for London and Paris, two most visited cities in the world, between December 2010 and February 2011. After dropping no-star hotels, our data set includes 665 London hotels and 769 Paris hotels. We collected following features for each hotel: the number of domestic and international bookings, average room price per night, star rating, the number of customer reviews, average customer rating and score breakdown (see Figure 3 for a sample hotel rating score). We calculate the service rating by taking the average of each dimension's score except location. Chain is a dummy variable taking the value of 1 if a hotel is part of a hotel chain and 0 if it operates independently. In our dataset, there are international and local hotel chains. Table 1, Table

2 and Table 3 present the description of key variables, the descriptive statistics and the correlation matrix for key variables respectively. The snapshot of data source is provided in Figure 1.

4.2. Empirical Model for Proportion of Domestic Bookings

Our variable of interest in this study is the ratio of domestic online bookings to total online bookings. We obtained the dependent variable by aggregating transaction level data. Since the dependent variable is a fraction, the conditional mean will be nonlinear and regression residuals will probably be heteroskedastic. Although ordinary least squares (OLS) can be used to explain a fractional variable, standard errors obtained through OLS method probably will be heteroskedastic and OLS estimators will be inefficient. Furthermore, the predicted values from an OLS regression can not be guaranteed to lie in the unit interval. Consequently, following prior research, we also consider a quasi-likelihood model with Bernoulli log-likelihood function [Papke and Wooldridge 1996, Kieschnick and McCullough 2003] and a parametric regression model with the beta distribution [Mittelhammer 1996, Kieschnick and McCullough 2003].

The linear regression model (OLS) is estimated as follows:

$$E\left(\left(\frac{\text{number of domestic bookings}}{\text{total bookings}}\right)_i \mid x_i\right) = \beta_0 + \beta_1 \ln(\text{number of customer reviews})_i + \beta_2 \ln(\text{average price})_i + \beta_3 \text{location rating}_i + \beta_4 \text{service rating}_i + \beta_5 \text{star}_i + \beta_6 \text{chain}_i + \beta_7 \ln(\text{rooms})_i \quad (1)$$

Since the numbers of customer reviews, average price and the number of rooms have relatively higher mean and standard deviation compared to other variables, we transform these variables by taking their natural logarithm. These transformations also help to increase the explanatory power of the models. We also estimate alternative models including quadratic terms of rating and star to above model. Since the coefficients of quadratic terms are not significant, we use the model specified in equation (1) using ordinary least squares (OLS). We include logarithmic transformation of number of hotel rooms as a control variable since we expect that hotels with higher number of hotel rooms are more likely to have higher number of reviews.

We also use following model proposed by Papke and Wooldridge [1996] to make sure that the predicted values of the dependent variable is in the unit interval.

$$E\left(\left(\frac{\text{number of domestic bookings}}{\text{total bookings}}\right)_i \mid x_i\right) = G\left(\begin{matrix} \beta_0 + \beta_1 \ln(\text{number of customer reviews})_i + \beta_2 \ln(\text{average price})_i \\ + \beta_3 \text{location rating}_i + \beta_4 \text{service rating}_i + \beta_5 \text{star}_i + \beta_6 \text{chain}_i + \beta_7 \ln(\text{rooms})_i \end{matrix}\right)$$

where $G(\cdot)$ is the logit function, i.e., $G(X\beta) = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$

In this model, we make an assumption about the functional form of the conditional expectation function but no distributional assumption. Model parameters are estimated using a quasi-likelihood method with the Bernoulli log-likelihood function,

$$L_i = y_i \log(G(X_i\beta)) + (1 - y_i) \log(1 - G(X_i\beta))$$

We estimate the model parameters by maximizing the log-likelihood function, $\max_{\beta} \sum_i L_i$. The model can also be estimated using a generalized linear model (GLM) with a logit link. GLM is a generalization of OLS. Furthermore, GLM can be used to estimate fractional and binary response variables and models with heteroskedacity and autocorrelation.

Lastly, we consider a parametric regression model with the beta distribution. Beta distribution is part of the exponential family which forms the basis for generalized linear model framework [Mittelhammer 1996]. Earlier studies found that beta distribution provide good fit for proportional data [Hviid and Villadsen 1995; Kieschnick and McCullough 2003]. Consequently, we use the following probability density function of the beta distribution in alternative parameterization which corresponds to the GLM convention,

$$f(y) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}$$

where Γ, μ and ϕ are the gamma function, the location parameter, and scale parameter respectively and

$E(y) = \mu$ and $Var(y) = \frac{\mu(1-\mu)}{1+\phi}$ [Paolino 2001]. For the estimation of parameters, we maximize the following

log-likelihood function, $\max_{\beta} \sum_i L_i$, where

$$L_i = \ln(\Gamma(\phi)) - \ln(\Gamma(\mu\phi)) - \ln(\Gamma((1-\mu)\phi)) + (\mu\phi - 1)\ln(y_i) + ((1-\mu)\phi - 1)\ln(y_i - 1)$$

4.3. Empirical Model for Online Hotel Bookings

We also investigate the drivers of online domestic and international bookings for the following reasons. First, we would like to understand the factors which may impact domestic and international bookings differently. Although the determinants of online hotel booking are investigated in earlier research such as Ögüt and Taş [2012], our dataset enable us to perform a more detailed analysis of online hotel bookings. Specifically, our work differs from Ögüt and Taş [2012] in the following aspects. First, prior research uses the number of reviews as proxy for number of sales data whereas we use actual sales data. Second, unlike prior studies, we have data for domestic and international bookings. Third, we analyze the impact of customer ratings in two main dimensions: location and service. Fourth, we calculate average room price over all transactions instead of using standard double room price [Ögüt and Taş 2012]. Therefore, we believe that our analysis will provide more reliable and detailed results. We use the following model in order to estimate domestic and international bookings:

$$\ln(\text{Sales})_i = \beta_0 + \beta_1 \ln(\text{average price})_i + \beta_2 \text{location rating}_i + \beta_3 \text{service rating}_i + \beta_4 \text{star}_i + \beta_5 \text{chain}_i + \beta_6 \ln(\text{rooms})_i$$

Our dependent variables are the number of domestic and international bookings and we used corresponding price for each dependent variable. We control for hotel size by including the number of rooms in our regression model as we expect a positive correlation between the number of bookings and hotel size. Similar to previously presented models for proportional data, we take the natural logarithm of sales, price and number of rooms, since the transformation increases the explanatory power of the model and allows us to compute elasticities of independent variables with respect to each dependent variable. We also use chain dummy in order to test whether a hotel with chain name impacts booking behavior of domestic and international customers differently. We estimate the model with OLS and Table 7 reports the regression results.

5. Estimation Results of Models

5.1. Estimation Results for Proportion of Domestic Booking

The estimation results of the models presented in Section 4.2 are provided in Table 4. Before discussing the results, there are a couple of comments worth making about computational issues. We checked whether there is heteroskedasticity in our data sets. Breusch-Pagan test rejects the null hypothesis of homoskedastic errors. In Table 4, we report the estimates of heteroskedasticity-robust standard errors. We also tested whether there is multicollinearity in our data sets using variance inflation factors (VIFs) score for each of the explanatory variables and we did not find multicollinearity problem in our data sets.

We also verify whether the distributional assumptions of above regression models are supported by data. A visual examination of data shows that the distribution of the proportion of domestic sales does not fit into normal distribution since the distribution of the data is skewed to the right. Both Jarque-Bera and Shapiro-Wilk tests reject the null hypothesis that residuals of the linear regression model follows normal distribution at 1% significance level. Furthermore, quantile-quantile plot of data versus a beta distribution suggests that data fits to the beta distribution well. Although the assumptions of the linear regression model are rejected, we cannot reject the parametric regression model with the beta distribution and the quasi-likelihood model in which no distributional assumption is made. Box-Cox test of the functional form is also performed for the conditional mean. We reject the null hypothesis of indicating that functional form of the conditional mean is non-linear.

We also provide Wald test statistics for the overall significance of models in Table 4. We find that all models are statistically significant at 1% significance. While R^2 serves as a goodness-of-fit measure with a clear definition in linear regression models, we do not have such a measure for non-linear models. Consequently, we use Pearson chi-square statistic to test the fit of the quasi-likelihood model against observed data. Since the test statistic was not significant, we could not support the hypothesis that data do not fit to the model. We also report Akaike's Information Criteria (AIC) statistics in Table 4. AIC is one of the widely used statistics for the selection of best model among the alternatives. Lower values of AIC statistic indicate the performance of the model is the best compared to other models. We observed that Beta regression model is superior to other models based on the comparison of AIC statistics and the visual inspection of data. We would like to note that although the AIC statistic of linear regression is the lowest with negative sign, this does not show that the linear regression model performs better than the quasi-likelihood model which has a positive AIC statistic. We already know that the linear model does not fit the data and it cannot ensure the predicted values of our fractional dependent variable falls inside the unit interval. Furthermore, the AIC statistic of linear regression is almost always negative when the dependent variable is a fraction as the likelihood in this case will always be larger than 1.

While the marginal effect of a change in one of the regressors on the expected conditional mean of the dependent variable is equal to the relevant coefficient estimate in the linear regression model, this is not the case in non-linear regression models as the functional form of the conditional mean of linear regression model is different from that of non-linear regression models. For this reason, we also provide the estimates of marginal effects evaluated at the sample averages of the regressors in Table 5.

As we noted before, we divided customer rating into service rating and location rating. To ensure that these two metrics summarize different dimensions of customer rating, we also performed factor analysis. We observed that location rating and other components of customer rating fall into different factors both in Paris and London datasets. This is evident from the correlations between service rating and location rating as well.

Our first hypothesis claims that there is an inverse relationship between hotel price and the proportion of domestic sales. We find that (the natural logarithm of) price is significant and negatively associated with proportion of domestic online sales in all model specifications both in Paris and London data sets. Specifically, a ten percent increase in average room price decreases the proportion of domestic sales by ranging from 0.0201 (0.0085) to 0.0216 (0.0093) unit for Paris (London) hotels. Thus, we find support for our first hypothesis.

Our second and third hypotheses investigate the effect of location rating, service rating and star rating on the proportion of domestic sales. Our empirical analysis shows that the coefficient of location rating is negative and significant both in Paris and London. Specifically, a one-unit increase in location rating decreases the proportion of domestic sales by ranging from 0.031 (0.053) to 0.035 (0.057) for Paris (London) hotels. Thus, we are able to confirm H3.a by showing that a higher location rating leads to a lower proportion of domestic sales. Furthermore, we find that the coefficient of service rating is positive and significant in Paris and London datasets. In particular, a one-unit increase in the service rating increases the proportion of domestic sales by up to 0.016 (0.032) for Paris (London) hotels. Thus, we can support H3.b by showing that a higher service rating leads to a higher proportion of domestic sales. Our analysis also provides support for the effect of star rating on the proportion of domestic sales. Our results show that the coefficient of star rating is positive and significant both for London and Paris hotels. Furthermore, a one-unit increase in star rating increases the proportion of domestic sales by up to 0.025 (0.025) points in Paris (London) datasets.

Our fourth hypothesis claims that (natural logarithm of) review volume is negatively associated with the proportion of domestic sales. Our empirical results confirm this claim both in Paris and London data sets. We also observe that a ten percent increase in review volume decreases the proportion of domestic sales by up to 0.0013 (0.0013) unit in Paris (London) datasets.

We also found that chain hotels have 0.082 units higher proportion of domestic customers compared to non-chain hotels in London. However, the coefficient of chain dummy is not significant for Paris hotels. Thus, hypothesis 5 is not supported. To shed light on the differing findings, we report average price of chain and non-chain hotels for Paris and London in Table 6. As it is seen from Table 6, chain hotels in London charge relatively higher prices to international customers compared to chain hotels in Paris. This observation might explain why chain hotels in London have higher proportion of domestic customers while this is not the case for Paris hotels.

5.2. Estimation Results for Online Hotel Bookings

We have the following observations from the estimation of hotel booking data. Since we use logarithmic transformation of sales and price in all models, the coefficient of natural log transformation of price are interpreted as elasticity. We found that coefficient of price is negative and significant at the 1% significance level for all specifications both for London and Paris hotels. Specifically, a 1% increase in price decreases domestic (international) sales by 2.08 % (1.14%) for hotels in Paris while 1% increase in price decreases domestic (international) sales by 1.77 % (1.72%) for hotels in London. Consequently, domestic customers are more sensitive to the price changes than international customers. This result supports our first hypothesis as well. For the sake of completeness, we also consider all customers and find that 1% increase in price decreases sales by 1.41% (1.84%) at Paris (London) hotels.

Our results also show that the coefficient of star rating is not significant in any of the model specifications for both Paris and London datasets implying that change in star rating does not affect online hotel bookings. This finding suggests that star rating is losing its importance as customers think that customer ratings and reviews provide more reliable information [Öğüt and Taş 2012]. Although star rating is not significant, we observe that the coefficient of star rating is higher in the regressions for domestic bookings.

Like star rating, service rating measures the quality of service offered while location rating measures the convenience of the environment that hotel operates in. We observe that location rating is positively associated with both domestic and international bookings. Furthermore, regression results suggest that location rating's impact on international bookings is higher. Moreover, service rating's impact on domestic bookings is higher than that on international bookings. Thus, domestic customers (international customers) are more sensitive to changes in service

rating (location rating). These observations are consistent with hypothesis 3.a. and 3.b and provide additional insight for the explanation of these hypotheses.

Even though we did not take log transformation of service and location rating, it is possible to calculate their elasticity with respect to sales at their respective means by multiplying their coefficients and means as the dependent variable is in natural logarithmic form. By using the coefficient estimates of Table 7 and mean values of independent variables from Table 2, Table 8 presents the elasticities of different independent variables with respect to sales for Paris and London datasets. We do not report the elasticity for variables if they are not significant. Results show that a 1% increases in location rating increase hotel booking by 1.42% (1.40%), 2.69% (4.49%) and 2.90% (3.62%) at the domestic, international and all customers segments of Paris (London) hotels respectively at the mean level of location rating. Moreover, a 1% increase in service rating increases sales ranging from 2.29% to 2.56% for hotels in Paris and 1.34% to 2.08% for hotels in London at the mean level of service rating.

We found that the while the coefficient of chain dummy for domestic customers is not significant, the coefficient of chain dummy for international customers is significant with negative sign. The possible reason for this is that the price premium charged to international customers by hotel chains is high compared to services offered. Thus, if a hotel is part of a hotel chain, it is less preferable from international customers. Since the proportion of international customers is higher than domestic customers in both datasets, we also find negative effect of chain name on hotel bookings when both domestic and international customers are considered together. As chain is a dummy variable, $100 * (e^{\text{coefficient-variance of coefficient}/2} - 1)$ gives unbiased estimate of the percentage change on the dependent variable [Kennedy 1981]. The statistics is computed as 65% and 83% for international bookings in Paris and London respectively. Hence, chain hotels decreases the number of international bookings by 35%(17%) compared to non-chain hotels for Paris (London) hotels.

6. Discussions and Conclusion

In this study, we analyze the effect of review rating, review volume and hotel characteristics on the proportion of online domestic bookings. We find that most of our empirical results are consistent with our expectations which form the basis of our hypotheses. Price, review volume and location rating are significant and negatively associated with proportion of domestic online sales in all model specifications. Furthermore, service and star rating are significant and positively associated with proportion of domestic online sales in all model specifications. However, chain name does not lead to higher proportion of international customers. We also report marginal effect of these variables on our dependent variables.

We have also estimated the number of domestic and international bookings. We find that price has a significant and negative impact while location and service rating have a significant and positive impact on the number of domestic and international bookings. However, the impact of star rating on the number of domestic and international bookings is not significant. The number of international bookings decreases while the number of domestic bookings is not affected by whether a hotel is a part of hotel chain. Furthermore, domestic customers are more sensitive to changes in price and service rating while international customers are more responsive to changes in location rating. These findings support our earlier findings and provide additional insights for the explanation of our hypotheses.

Our results suggest that that domestic and international customers' online hotel booking behaviors show significant differences. Thus, different marketing strategies should be designed for domestic and international customers in the hotel booking process. Furthermore, our empirical models allow hotel managers to estimate proportion of domestic (international) sales based on publicly available information retrieved from the Internet. Moreover, models presented in this study can be used to develop techniques for a preferred mix of domestic and international sales, make an efficient use of fixed capacities, improve hotels' occupancy rates and generate more revenue and profits.

REFERENCES

- Akerlof, G.A. "The market for 'lemons': Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, Vol. 84, No.3:488-500, 1970.
- Ariffin, A.A.M. A. Maghzi, and A. Norzalita, "Understanding Hotel Hospitality and Differences between Local and Foreign Guests," *International Review of Business Research Papers*, Vol.7, No. 1:340-349, 2011.
- Ba, S. and P. Pavlou, "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly*, Vol. 26, No. 3:243-268, 2002.
- Bharadwaj, S. G. R. P. Varadarajan, and J. Fahy, "Sustainable Competitive Advantage in Service Industries: A Conceptual Model and Research Propositions," *Journal of Marketing*, Vol. 57:83-99, 1993.
- Callan, R. "Attributional Analysis of Customer's hotel Selection Criteria by U.K. Grading Scheme Theories," *Journal of Travel Research*, Vol. 36, No. 3:20-34, 1998.

- Cantalops, A.S. and F. Salvi, "New consumer behavior: A review of research on eWOM and hotels," *International Journal of Hospitality Management*, Vol. 36:41–51, 2014.
- Chevalier, J. A. and D. Mayzlin, "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, Vol.43, No:3:345–354, 2006.
- comScore and the Kelsey Group, "Online Consumer-Generated Reviews Have Significant Impact on Offline Purchase Behavior," 2007.
http://www.comscore.com/Press_Events/Press_Releases/2007/11/Online_Consumer_Reviews_Impact_Offline_Purchasing_Behavior, Accessed 10.05.2014.
- Duan, W. B. Gu, and A.B. Whinston, "Do online reviews matter? — An empirical investigation of panel data," *Decision Support Systems*, Vol. 45, No. 4:1007-1016, 2008.
- Forman, C. A. Ghose, and B. Wiesenfeld, "Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Discloser in Electronic Markets," *Information Systems Research*, Vol.19, No.3:292-313, 2008.
- Ghose, A. P.G. Ipeiritos, and B. Li, "Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowd-Sourced Content," *Marketing Science*, Vol. 31, No. 3:493-520, 2012.
- Google/IPSOS OTX, "MediaTraveler's Road to Decision 2011,"
http://www.gstatic.com/ads/research/en/2011_TravelersRoadtoDecision2011.pdf, 2011. Accessed 10.05.2014.
- Gu, B. and Q. Ye, "First Step in Social Media- Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management*, Vol. 23, No.4:570–582, 2014.
- Hviid, M. and B. Villadsen, "Beta Distributed Market Shares in a Spatial Model With an Application to the Market for Audit Services," *Review of Industrial Organization*, Vol. 10:737-747, 1995.
- Kennedy, P. "Estimation with Correctly Interpreted Dummy Variables in Semi Logarithmic Equations," *American Economic Review*, Vol. 71:801, 1981.
- Kieschnick, R., and B.D. McCullough, "Regression analysis of variates observed on (0, 1): percentages, proportions and fractions", *Statistical Modelling*, Vol. 3, No. 3:193-213, 2003.
- Marcussen, C. "Trends in European Internet Distribution - of Travel and Tourism Services," <
<http://www.crt.dk/UK/staff/chm/trends.htm>, 2008, Accessed 23.04.2014.
- Mauria, A.G. and R. Minazzi, "Web reviews influence on expectations and purchasing intentions of hotel potential customers", *International Journal of Hospitality Management*, Vol. 34:99–107. 2013.
- Mittelhammer R.C. "*Mathematical statistics for economics and business*," New York: Springer-Verlag, 1996.
- O'Neill, J. and Q. Xiao, "The Role of Brand Affiliation in Hotel Market Value," *Cornell Hospitality Quarterly*, Vol. 47, No. 3:210-223, 2006.
- Öğüt, H. and B.K.O. Taş, "The influence of internet customer reviews on the online sales and prices in hotel industry," *The Service Industries Journal*, Vol. 32, No. 2:197-214, 2012.
- Paolino, P. "Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables," *Political Analysis*, Vol. 9, No. 4, 325-346, 2001.
- Papke, L. and J. Wooldridge, "Econometric methods for fractional response variables with an application to 401(K) plan participation rates," *Journal of Applied Econometrics*, Vol. 11: 619–632, 1996.
- Park, D.H. J. Lee, and J. Han, "The effect of online consumer reviews on consumer purchasing intention: the moderating role of involvement," *International Journal of Electronic Commerce*, Vol. 11, No. 4:125–148, 2007.
- Pathak, B., R.S. Garfinkel, R.D. Gopal, R. Venkatesan, and F. Yin, "Empirical Analysis of the Impact of Recommender Systems on Sales," *Journal of Management Information Systems*, Vol. 27, No.2:159-188, 2010.
- PhoCusWright, "Global online travel overview second edition report", April 2011.
- Spence, A.M. "Job market signaling," *Quarterly Journal of Economics*, Vol. 87, No. 3:355–374, 1973.
- Vermeulen, I. E. and D. Seegers, "Tried and tested: The impact of online hotel reviews on consumer consideration," *Tourism Management*, Vol. 30:123-127, 2009.
- Yüksel, A. "Shopping experience evaluation: a case of domestic and international visitors," *Tourism Management*, Vol. 24:751–759, 2004.
- Yesawich, P. C. "So Many Brands, So Little Time", *Lodging Hospitality*, Vol.52, No. 9:16, 1996.
- Xie, H., L. Miao, B.Y. Lee, "Consumers' responses to ambivalent online hotel reviews: The role of perceived source credibility and pre-decisional disposition," *International Journal of Hospitality Management*, Vol. 30: 178–183, 2011.
- Zhang, Z., Q. Ye, R. Law, and Y. Li, "The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews," *International Journal of Hospitality Management*, Vol. 29, No. 4:694–700, 2010.

APPENDIX

Table 1. Description of variables

Variable	Description and measure
prdombook	Proportion of domestic online bookings
price	Average price per room per night over all transactions between December 2010 and February 2011
service rating	Hotel's average review rating excluding location rating
location rating	Hotel's average location rating
nreview	The number of hotel's customer reviews
Star	Hotel's star rating
rooms	The number of hotel rooms
chain	Dummy variable taking the value of 1 if a hotel is a part of hotel chain and 0 if hotel operates independently.

Table 2.a Descriptive statistics for Paris data set

	Mean	Standard Deviation	Minimum	Maximum
ln(nreview)	4.882	1.019	1.609	7.748
ln(price)	4.803	0.437	3.890	6.507
location rating	8.518	0.772	5.50	9.80
service rating	7.260	0.850	4.02	9.46
Star	2.853	0.780	1	5
ln(rooms)	3.739	0.637	2.079	6.932
chain	0.317	0.317	0	1
<i>N</i>	769			

Table 2.b Descriptive statistics for London data set

	Mean	Standard Deviation	Minimum	Maximum
ln(nreview)	5.130	1.157	1.609	8.732
ln(price)	4.699	0.457	3.406	6.055
location rating	8.425	0.841	4.8	9.8
service rating	7.169	1.063	3.72	9.56
star	3.327	1.063	1	5
ln(rooms)	4.103	1.049	1.609	6.960
chain	0.398	0.489	0	1
<i>N</i>	665			

Table 3.a Correlation matrix of key variables for Paris data set

	prdombook	ln(nreview)	ln(price)	location rating	service rating	star	ln(rooms)	chain
prdombook	1.00							
ln(nreview)	0.04	1.00						
ln(price)	-0.48***	-0.26***	1.00					
location rating	-0.41***	-0.04	0.55***	1.00				
service rating	-0.23***	0.10**	0.56***	0.30***	1.00			
star	-0.31***	-0.22***	0.79***	0.31***	0.42***	1.00		
ln(rooms)	-0.08**	0.15***	0.23***	-0.03*	0.16***	0.27***	1.00	
chain	-0.01	-0.01	0.17***	-0.06*	0.17***	0.24***	0.42***	1.00

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.b Correlation matrix of key variables for London data set

	prdombook	ln(nreview)	ln(price)	location rating	service rating	star	ln(rooms)	chain
prdombook	1.00							
ln(nreview)	-0.02	1.00						
ln(price)	0.01	-0.31***	1.00					
location rating	-0.27***	0.00	0.56***	1.00				
service rating	0.23***	-0.06	0.63***	0.24***	1.00			
star	0.15***	-0.22***	0.78***	0.35***	0.62***	1.00		
ln(rooms)	0.27***	0.22***	0.37***	0.19***	0.34***	0.40***	1.00	
chain	0.35***	-0.03	0.39***	0.18***	0.41***	0.41***	0.59***	1.00

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.a Results for different regression models of proportion of domestic sales for Paris data set

	(OLS)	(Q-LM)	(Beta)
ln(nreview)	-0.013** (0.005)	-0.060** (0.026)	-0.049* (0.025)
ln(price)	-0.201*** (0.027)	-0.101*** (0.130)	-0.946*** (0.125)
location rating	-0.035*** (0.009)	-0.154*** (0.042)	-0.144*** (0.041)
service rating	0.015* (0.008)	0.077** (0.036)	0.074** (0.036)
star	0.024** (0.012)	0.118** (0.056)	0.110** (0.054)
ln(rooms)	-0.003 (0.008)	-0.020 (0.045)	-0.018 (0.042)
chain	0.011 (0.011)	0.058 (0.055)	0.060 (0.051)
constant	1.478*** (0.083)	4.836*** (0.401)	4.426*** (0.386)
AIC	-895.153	664.488	-999.309
r2	0.282		
Wald	44.2	289.259	281.720
Ll	455.576	-324.244	508.654
N	796	769	769

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.b Results for different regression models of proportion of domestic sales for London data set

	(OLS)	(Q-LM)	(Beta)
ln(nreview)	-0.012** (0.005)	-0.060** (0.025)	-0.048** (0.024)
ln(price)	-0.088*** (0.023)	-0.429*** (0.111)	-0.391** (0.105)
location rating	-0.057*** (0.007)	-0.255*** (0.034)	-0.246*** (0.033)
service rating	0.031*** (0.007)	0.150*** (0.033)	0.134*** (0.031)
star	0.022** (0.009)	0.116** (0.045)	0.098** (0.044)
ln(rooms)	0.025*** (0.006)	0.118*** (0.030)	0.114*** (0.029)
chain	0.082*** (0.014)	0.376*** (0.068)	0.374** (0.067)
constant	0.843*** (0.083)	1.616*** (0.383)	1.506** (0.371)
AIC	-799.666	579.744	-871.834
r2	0.297		
Wald	41.83	301.853	291.260
Ll	407.833	-281.872	444.917
N	665	665	665

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.a Marginal effects evaluated at the sample mean of the regressors for Paris data set

	OLS	(Q-LM)	(Beta)
ln(nreview)	-0.013	-0.013	-0.010
ln(price)	-0.201	-0.216	-0.203
location rating	-0.035	-0.033	-0.031
service rating	0.015	0.016	0.016
star	0.024	0.025	0.023
ln(rooms)	-0.003	-0.004	-0.003
chain	0.011	0.012	0.013

Table 5.b Marginal effects evaluated at the sample mean of the regressors for London data set

	OLS	(Q-LM)	(Beta)
ln(nreview)	-0.012	-0.013	-0.010
ln(price)	-0.088	-0.093	-0.085
location rating	-0.057	-0.055	-0.053
service rating	0.031	0.032	0.029
star	0.022	0.025	0.021
ln(rooms)	0.025	0.025	0.025
chain	0.082	0.082	0.082

Table 6.a. Average Prices of Chain and Non-chain Hotels in Paris

Paris	Number of Hotels	Average Price for Domestic Customer	Average Price for International Customer
non-chain hotels	525	119.29 €	129.71 €
chain hotels	244	144.63 €	153.23 €
difference		25.34 €	23.52 €

Table 6.b. Average Prices of Chain and Non-chain Hotels in London

London	Number of Hotels	Average Price for Domestic Customer	Average Price for International Customer
non-chain hotels	400	98.43 £	106.61 £
chain hotels	265	140.94 £	151.15 £
difference		42.51 £	44.54 £

Table 7.a OLS regression for sales in Paris data set

	Domestic	International	All
ln(price)	-2.089*** (0.156)	-1.143*** (0.173)	-1.417*** (0.155)
location rating	0.167*** (0.054)	0.316*** (0.054)	0.341*** (0.052)
service rating	0.353*** (0.047)	0.316*** (0.054)	0.332*** (0.048)
Star	0.086 (0.069)	-0.041 (0.072)	-0.005 (0.064)
ln(rooms)	0.479*** (0.054)	0.500*** (0.056)	0.489*** (0.051)
chain	-0.084 (0.079)	-0.180** (0.079)	-0.153** (0.074)
constant	7.598*** (0.507)	2.622*** (0.552)	4.696*** (0.490)
R2	0.367	0.189	0.252
N	769	769	769

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7.b OLS regression for sales in London data set

	Domestic	International	All
ln(price)	-1.773*** (0.156)	-1.726*** (0.160)	-1.841*** (0.153)
location rating	0.167*** (0.054)	0.534*** (0.058)	0.430*** (0.055)
service rating	0.291*** (0.048)	0.187*** (0.047)	0.242*** (0.044)
star	0.010 (0.067)	-0.062 (0.066)	-0.025 (0.061)
ln(rooms)	0.573*** (0.048)	0.474*** (0.046)	0.504*** (0.044)
chain	0.043 (0.108)	-0.403*** (0.112)	-0.263** (0.104)
constant	6.052*** (0.520)	5.285*** (0.540)	6.406*** (0.494)
R2	0.352	0.315	0.349
N	665	665	665

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$


Table 8.a Elasticities with respect to Sales in Paris

	Domestic	International	All
price	-2.089	-1.143	-1.417
location rating	1.422	2.691	2.904
service rating	2.562	2.294	2.410
star	-	-	-
chain	-	0.830	0.853
rooms	0.479	0.500	0.489

Table 8.b Elasticities with respect to Sales in London

	Domestic	International	All
price	-1.773	-1.726	-1.841
location rating	1.406	4.498	3.622
service rating	2.086	1.340	1.734
star	-	-	-
chain	-	0.659	0.760
rooms	0.558	0.365	0.425

Recently booked in Paris



Buffault Grand Hotel Lafayette ★★★★★

6, Rue Buffault, 09. Opéra, Paris [Show map](#)


Score from 438 reviews: *Pleasant*, **6.3** 🗨️

Single Room booked today at 14:22 from Ukraine for € 234 (3 nights)

from

€ 72

Situated only 3 minutes from the Opera, the major shops, and the business centre, (insurances companies, salle Drouot), the entirely renovated Hotel Lafayette Buffault offers you the comfort of its 45... [More...](#)



Libertel Montmartre Duperre ★★★★★

32 Rue Duperre, 09. Opéra, Paris [Show map](#)

Score from 1055 reviews: *Good*, **7.9** 🗨️

Comfort Room (Special Offer) booked today at 14:22 from France for € 64 (1 night)

There is 1 user browsing this hotel's page right now.

from

€ 45

Libertel Montmartre Duperre offers free Wi-Fi and excellent travel links. It is situated in a lively part of Paris, near the Moulin Rouge, the Sacre-Coeur and Montmartre. [More...](#)



Napoleon Paris ★★★★★

40, Avenue de Friedland, 08. Champs Elysées, Paris [Show map](#)

Score from 283 reviews: *Very good*, **8.1** 🗨️

Suite booked today at 14:22 from Lebanon for € 372 (1 night)

There is 1 user browsing this hotel's page right now.

from

€ 189

Figure 1: Snapshot of the Data Source

How did you like the hotel on these points...?



Figure 2: Customer Satisfaction Survey Sent to Customers after Their Stay

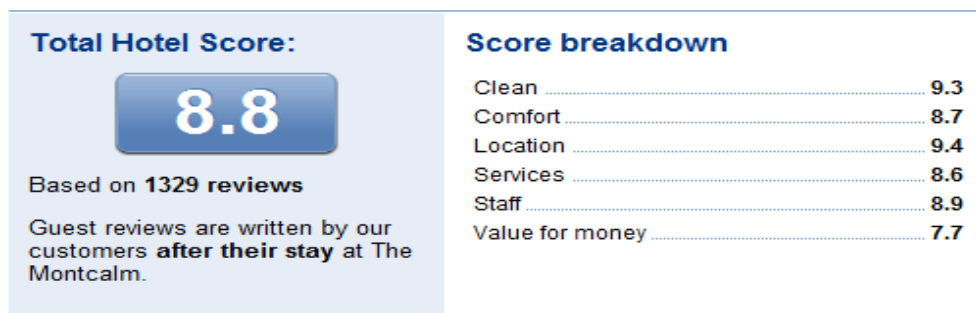


Figure 3: A Sample Total Hotel Score and Score breakdown