Universitat de Lleida

Hotels that most rely on Booking.com

Eva Martin-Fuentes – University of Lleida

Juan Pedro Mellinas – Universidad Internacional de la Rioja

Purpose

The purpose of this article is to know which hotels most rely on Booking.com, investigating the level of presence on Booking.com through the number of reviews per room by country, hotel size, hotel category, and managerial form. Neither the company nor the hotels provide this information, so we use the number of reviews as an indicator of estimated sales.

Design/methodology/approach

Data from 33,996 hotels worldwide is downloaded from Booking.com using a web browser automatically controlled, developed in Python, that simulated a user navigation (clicks and selections). The comparison between independent hotels and hotels belonging to a chain is performed by a Student's-t distribution test and the comparison of hotel categories and hotel size is analyzed by a one-way ANOVA test.

Findings

The results show that three factors clearly influence the use level of Booking.com: independent vs. chain hotels, small vs. large hotels, and low vs. high category hotels worldwide. We observe also that hotels from Europe are the ones that more rely on Booking.com.

Originality/Value

The originality of this research is to identify the factors that make hotels to have a greater (lesser) dependence on Booking.com within each destination and geographical area. Moreover, the use of big data from hotels worldwide allows us to know the level of use of Booking.com in dozens of countries, especially those with the highest tourist activity.

This work expands the capabilities of big data in the hospitality industry research and, with a simple ratio, this study counteracts the lack of public data of hotel sales through Booking.com. This new approach could be extended to the analysis of other OTAs, which use similar reviews systems.

1. Introduction

The emergence of the Internet since the 90s has been transforming the way that products and services are distributed to customers around the world (Berne et al.,

2012). The hotel sector has been one of the most affected by this new way of sell its services, becoming an essential channel in the selling strategy of any accommodation business in the world.

In this study we focus on investigating the level of presence of Booking.com, the main online travel agency (OTA) for the commercialization of hotels worldwide, in different countries and types of hotels. We can easily know the number of establishments that this website offers in each destination, but we know almost nothing about the sales of these establishments using Booking.com, neither in absolute numbers, nor as a percentage of their distribution mix. This data is especially difficult to obtain from the hotels or from Booking.com, since it is very sensitive economic information.

To estimate the number of reservations made through Booking.com in a hotel, we based on the number of reviews registered in each establishment. After their trip, guests receive an email (Figure 1) inviting them to write about their stay (Booking, 2017a). This is the only way to leave a review on Booking.com.

Insert Figure 1

We must consider that not all guests who receive the Booking.com email inviting to write a review, collaborate with the company, but we assume that customers have a constant probability of posting a review on Booking.com. It means that the number of reviews is expected to be closely correlated to the number of room sales. Therefore, the number of published reviews on Booking.com might be taken as a proxy for sales of a hotel's rooms through this travel agent.

This methodology has been used before, using Amazon.com's sales rankings to infer product sales (Chevalier and Mayzlin, 2006; Ghose and Ipeirotis, 2006),others focusing in the relationship between the number of reviews of online games and the sales of online games (Zhu and Zhang, 2010) or the relationship between reviews and book sales (Forman et al., 2008).

In the hotel industry, a research with data from Booking.com, assumed that the number of hotel reviews is a constant fraction of sales for all the hotels (Öğüt and Onur Taş, 2012). (Ye et al., 2008) considered the number of online reviews as a proxy for hotel sales (in Ctrip.com and Elong.net) and validated their approach with a 19 hotels sample. Results showed a significant positive relationship between the number of reviews and the number of hotel bookings with an R-square of 68%. These authors wrote a new paper assuming this relationship and validating it again with a 27 hotels sample for Ctrip.com. In this case they got a R-square value of 44.1%(Ye et al., 2011). These three papers have had some popularity among the scientific community, with 148, 32 and 589 citations respectively, according to Google scholar.

This relationship between hotel reviews and sales has only been tested in two small samples of Chinese hotels, so its accuracy and validity has been questioned (Schuckert et al., 2015). However, several authors have developed their research based on its validity as a reliable proxy (Lu et al., 2012, 2014; Viglia et al., 2014). Recently, this methodology was used to determine the level of dependency of Spanish hotels on Booking.com, establishing a relationship between the number of

reviews collected in a year and the number of rooms in each establishment, defining a ratio called Reviews per Room (RpR) (Mellinas, 2017).

We use this ratio as a proxy of the level of use of Booking.com by hotels: a greater RpR will indicate a greater percentage of reservations through this website. If a hotel has twice the RpR than another it will be because its percentage of sales through Booking.com is double that in the other.

Using this ratio, we carry out a worldwide study on Booking.com. First of all, we identify the level of importance of Booking.com in dozens of countries around the world, especially those with the highest tourist activity worldwide. On the other hand, we identify the factors that make hotels, individually, have a greater or lesser dependence on Booking.com within each destination and geographical area. The results obtained highlight different profiles of hotels that have a greater tendency to use Booking.com.

Booking.com

There are two companies that dominate hotel distribution in Internet worldwide. On one hand, Expedia (Expedia, Hotels.com, Orbitz, Travelocity, Hotwire, Wotif, etc.) and in the other, Priceline, which uses six primary brands: Booking.com, priceline.com, agoda.com, KAYAK, Rentalcars.com and OpenTable (Priceline, 2017). Booking.com accounts for more than two-thirds of total revenue in the group (Levy, 2013) and is the world leader in booking accommodations online. It is the leader OTA in most European countries with 62% of global market share (Schegg, 2016). The website is available in 43 languages, offers over 1,300,000 accommodations and covers over 100,000 destinations in 220 countries and territories worldwide.

Booking.com is also a huge reviews database that features over 140 million verified reviews written by guests after their stay. These reviews have a numerical part in a scale of 2.5-10 (Mellinas et al., 2015) and a qualitative part referring to the text that users voluntarily write. Booking.com has got a clear leadership position in the number of reviews, when we compare it with other OTAs, hosting 39 percent of all reviews worldwide (Murphy, 2017) and probably a higher percentage if we focus in the European market.

Although Booking.com is based in Amsterdam (Priceline bought the company in 2005) and was primarily focused in the European market, in recent years it has faced an expansion worldwide, for example in the United States (Booking.com, 2013). In this market it enters into an apparent direct competition with Priceline, which is focused on the North American market. But both companies seems to operate in an autonomous way; Booking.com is the world's top hotel-reservation service, while Priceline is the leading provider of discount-travel services in the U.S. (Levy, 2013).

2. Literature Review

Academic researchers have been focusing on online distribution as alternative to increase reservations, revenue and profit (Pal and Mishra, 2017; Talluri and Van

Ryzin, 2006). Some players in the distribution network, gain power, which may lead to dependencies, particularly in OTAs, especially those which belongs to the cited Expedia and Priceline groups (Pearce, 2008). OTAs has had a profound effect in the way that hospitality industry works (Buhalis and Law, 2008), offering a vast amount of opportunities and alternative distribution systems (Toh et al., 2011). Although both traditional travel agencies and OTAs are equally important at present and have similar shares, the role of the former will inevitably decline in the coming years (Law et al., 2015).

OTAs helps hotels to sell their services, but reducing incomes because of commissions (Carroll and Siguaw, 2003; Ford et al., 2012). The average commission charged by Booking.com or Expedia is more than 18 per cent of the price (Garrigos-Simon et al., 2017) and hotels have some other costs with virtual operators: direct fraud, no shows, cancellations or non-conversion rates. It was assumed that the importance of intermediaries would decrease (Bennett and Lai, 2005; Tse, 2003) and direct channels (hotel website, social media, hotel App) would increase its market share (Ip et al., 2011), mainly because of the commission charged by OTAs. However direct bookings have reduced, while OTAs share has been increasing in levels around 25% (Schegg, 2016) or 40% (European Commission, 2017) in the European market.

These research from the European trade association of hotels (Schegg, 2016), showed that one player (Booking.com) control closely two thirds of OTA market, which means a Booking.com market share of around 15%. Additional data about "*OTA market shares by segment*" are provided by the cited research, showing figures of 23.4% for independent hotels and 19.9% for hotel chains, Similar differences between independent hotels and chain hotels can be observed in the report published by the European Union competition authorities (European Commission, 2017). Also variations by hotel category, from 33.5% to 17.0% and additional differences by hotel size, from 27.5% to 17.6%.

During the last years, hotel reviews databases have gained great importance, generating a large number of publications on this topic, as a valuable source of information for academic researchers and hoteliers (Cantallops and Salvi, 2014; Kwok et al., 2017). Whenever applicable, researchers are replacing the data sets collected through questionnaires and interviews by those collected from online services, with Booking.com and TripAdvisor being the most prominent sources (Stanisic, 2016). The increased use of big data collected from online review websites for research purposes is supported by automatically controlled systems, which acquire information about millions of reviews from thousands of hotels (Radojevic et al., 2015a) quickly, cheaply and conveniently. Once this huge databases are obtained, researchers have attempted to analyze and understand online traveler reviews by sophisticated technologies (Govers and Go, 2004; Ye, Law, et al., 2009; Ye, Zhang, et al., 2009).

We could consider Booking.com as a relevant example of crowdsourcing or crowdvoting for the hospitality industry (Garrigos-Simon et al., 2017). Crowdsourcing can be used to improve the design and installations of hotels, to reduce product defects and room maintenance issues in hotels. Moreover, hotels can develop new models of product and service excellence, personalization and authenticity (Richard et al., 2016). Mellinas et al. (2015) identified 12 papers using Booking.com reviews

database as source of information. Since then, papers have not stopped appearing academic papers dealing with hospitality issues and using Booking.com hotel reviews (Abrate and Viglia, 2016; Borges et al., 2015; Díaz and Rodríguez, 2017; Mariani and Borghi, 2018; Martin-Fuentes et al., 2018; Mellinas et al., 2016; Radojevic et al., 2015b, 2017).

Motivations behind writing and sharing online reviews have been discussed in the academic literature (Hennig-Thurau et al., 2004; Schuckert et al., 2015). Some authors argue that contributors are mostly driven by intrinsic and positive motives such as enjoyment, positive self-enhancement, concerns for other consumers or wanting to help the company rather than vengeance and the need to vent (Yoo and Gretzel, 2008). This altruistic phenomenon is considered an online reproduction of similar phenomena of friendliness and collaboration that exist in the offline or "real world" (Resnick and Zeckhauser, 2002).

Products and services reviews on the Internet show the best and the worst of people (Whitty and Joinson, 2008). On the one hand, anonymity favors users to give more honest opinions. But on the other, that anonymity, encourages some users to lie more than they would in real life. They are a way to show complaints after an unsatisfactory hotel experience (Chiappa and Dall'Aglio, 2012). The motivation for posting negative reviews ranges from taking revenge to warning others (Wetzer et al., 2007).

This could suggest reviewers much more likely to post reviews if they are extremely satisfied or disappointed. Previous research using Booking.com showed a very small percentage of travelers extremely disappointed. After analyze 662,991 reviews Booking.com scores were infrequently low, with more than 80% of review scores being higher than 6.0 and only 2% under 4 (Bjørkelund et al., 2012).

3. Methodology

Data for this study was collected using a web browser automatically controlled, developed in Python, that simulated a user navigation (clicks and selections) for Booking.com.

We automatically gathered number of reviews, ranking and scoring, hotel name, city, country, hotel category, number of rooms, if the hotel belongs to a chain or if it is independent, and date from which hotels start to work with Booking.com, corresponding to the top destinations in the world according to TripAdvisor Ranking. The ranking was downloaded by the ratings of "All reviewers" option, and we have only hotels into account, discarding other options such as apartments, villas, or hostels.

Having obtained 40,580 hotels on Booking.com, we discard those that were in Booking.com for less than 12 months in order to avoid possible bias, so we analyze a total of 33,996 hotels, as shown in Table 1.

We divide the data set into regions: America (AME), Asia and Pacific (ASP), Europe (EUR), and Middle East and Africa (MEA), as suggested in other researches (Banerjee and Chua, 2016; Martin-Fuentes, 2016).

Table 1. Data sample

Booking.com	
Countries	66
Destinations	441
Hotels	33,996
Total reviews	17,854,988
Min. Review	5
Max. Review	18,115

Source: Compiled by the authors based on data from Booking.com

The statistical calculations were performed using SPSS Software, version 20.

To quantify the reviews per room (RpR) we use a simple index obtained by dividing the number of reviews on Booking.com for a given hotel by its number of rooms (Mellinas, 2017). As Booking.com only keep in its website the reviews performed by users for the last 24 months, the RpR was divided by 2 to obtain the RpR per year. Those hotels that were on Booking.com for less than 24 months, the months on Booking.com were transformed to years (i.e. 18 months → 1.5 years) and the RpR was divided into the years that a hotel is on Booking.com.

As we were working with a large volume of data, applying the central limit theorem, the population of sample means was assumed to be normal.

To compare the RpR by independent hotels or hotels belonging to a chain, we performe a Student's-t distribution test.

To compare the RpR by hotel categories, the one-way ANOVA test is performed to determine whether there are any significant differences between the RpR mean on Booking.com according to the hotel category (1 to 5 stars). It tests the null hypothesis:

- H_0: $\mu_1 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ being μ_n the RpR mean by hotel category (n = 1, 2, 3, 4, 5).
- H_A: There are two or more group means that are significantly different from each other.

The comparative of the RpR mean by hotel size, a one-way ANOVA test is performed in three groups: small (from 1 to 20 rooms), medium (from 21 to 80 rooms), and large (more than 80 rooms).

4. Results

The results are presented to confirm or reject the comparative of RpR mean on Booking.com according to the managerial form (independent hotel or chain), the hotel category, the hotel size, and the region where the hotel is located.

An additional analysis, looking for a possible relationship between RpR and hotel scores obtained on Booking.com was also performed. The results offered a low Pearson correlation coefficient (ρ = .15), the highest correlation coefficient was in ASP (ρ = .28) and the lowest in EUR (ρ = .02).

4.1. RpR by Independent hotels vs chain hotels

To check the mean comparative for RpR between independent hotels and hotels belonging to a chain, a Student's t-test was performed with independent samples of variables, as the hotels were not the same.

The results allowed us to reject the null hypothesis that the RpR means were equal on both type of hotels, so we confirmed that RpR mean in independent hotels was higher than on those belonging to chains (t = 1.99, df = 33893, p < .001). Table 2 shows a highest RpR mean for the independent hotels.

Table 2. Descriptive statistics of RpR between independent hotels and chain hotels

	N	Mean	SD	Min	Max
Independent	27,322	5.47	7.06	.01	86.22
Chain	6,573	3.49	4.18	.01	40.89

Source: Compiled by the authors based on data from Booking.com

By regions, we confirm that independent hotels have a higher RpR mean than those belonging to a chain; the results in all cases were statistically significant (p < .001). The highest RpR mean difference is shown in hotels from Europe (t = 2.7. df = 15,595. p < .001), which confirm that independent hotels receive more guests from Booking.com than hotels belonging to a chain, as can be observed in Table 3, those surely have less dependency to OTAs.

Table 3. Descriptive statistics of RpR by regions

		N	Mean	SD	Min	Max
AME	Independent	5,272	4.03	5.11	.02	78.46
	Chain	1,729	2.13	2.69	.01	37.95
ASP	Independent	8,364	2.71	4.37	.01	69.67
	Chain	1,472	1.61	2.14	.01	21.83
EUR	Independent	12,619	8.08	8.35	.01	86.22
	Chain	2,978	5.38	4.98	.02	49.89
MEA	Independent	1,067	3.40	4.38	.01	40.63
	Chain	394	2.13	2.15	.05	13.61

Source: Compiled by the authors based on data from Booking.com

4.2. RpR by hotel category

To check the equality of variances, i.e., homoscedasticity, Levene's test was performed and the assumption of homogeneity was not met, because ($F_{(4, 29,811)}$ = 645.47, p < .001). Since the assumption of homogeneity of variance was not met for this data, we used the obtained Welch's adjusted F ratio ($F_{(4, 5155.48)}$ = 627.31,

p < .001). We can conclude that at least two of the five hotel categories groups differ significantly in their average review per room.

Then, analyzing in depth each hotel category determined by the hotel stars (from 1 to 5), the mean comparative was performed by an ANOVA test post hoc Games-Howell (not homogeneity of variance and not the same groups sizes) and the results show that the higher the category is, the lower RpR gets. All results are statistically significant, except RpR of hotels of 1 and 2 stars (p = 0.07), which confirm in these two groups (1-star and 2-star hotels) that there is no significant difference.

Table 4. Descriptive statistics of RpR by hotel category

RpR	N	Mean	SD	Min	Max
1-star	892	7.52	9.76	.06	83.50
2-stars	4,270	6.35	8.33	.02	86.22
3-stars	12,795	5.42	6.73	.01	69.67
4-stars	8,974	4.12	4.75	.01	47.38
5-stars	2,885	2.10	2.64	.01	30.88
Total	29,816	4.90	6.42	.01	86.22

Source: Compiled by the authors based on data from Booking.com

The results confirm that hotels with highest categories sell fewer rooms through Booking.com, otherwise, hotels with lowest categories have bigger dependency on Booking.com because they have more reviews, as shown in Table 4. Except between 1-star and 2-stars hotels that there is no mean difference.

4.3.   RpR by hotel size

The rooms are grouped into three categories in order to classify the hotel size: small (from 1 to 20 rooms), medium (from 21 to 80 rooms) and large (more than 80 rooms).

To check the equality of variances, i.e., homoscedasticity, Levene's test was performed and the assumption of homogeneity was not met, because (F(2, 33892) = 2816.69, *p* < .001). Since the assumption of homogeneity of variance was not met for this data, we used the obtained Welch's adjusted F ratio (F(2, 15176.69) = 2434.02, *p* < .001). We can conclude that at least two of the three hotel sizes groups differ significantly in their average review per room.

Then, analyzing in depth each hotel category determined by the hotel size, the mean comparative was performed by an ANOVA test post hoc Games-Howell (not homogeneity of variance and not the same groups sizes) and the results show that the higher the category is, the lower RpR gets as can be observed in Table 5. All results are statistically significant.

An one-way ANOVA Post hoc Games-Howell test was performed and the results show that the smaller the hotels are, the higher RpR has, showing that in all cases the difference is statistically significant (*p* < .001).

Table 5. Descriptive statistics of RpR by hotel size

RpR	N	Mean	SD	Min	Max
1-20	6,976	8.96	9.40	.13	86.22
21-80	15,017	5.34	6.20	.03	51.83
>80	11,902	2.50	3.16	.01	32.46
Total	33,895	5.09	6.65	.01	86.22

Source: Compiled by the authors based on data from Booking.com

By regions there are also significant differences in RpR means, with an ANOVA test taking as a group each region the regions with a higher RpR are EUR, AME, MEA and ASP.

The results by regions show the same pattern, the smaller the hotel is, the higher RpR have, as can be seen in Table X with the ANOVA test, showing all the results statistically significant ($p < .001$).

Table 6. Descriptive statistics of RpR by hotel size and by regions

Classification		N	Mean	SD	Min	Max
AME	1-20	1,510	6.54	7.04	.13	78.46
	21-80	2,564	3.63	4.09	.04	37.95
	>80	2,927	1.96	2.34	.01	23.79
	Total	7,001	3.56	4.70	.01	78.46
ASP	1-20	1,523	5.28	6.93	.13	69.67
	21-80	4,157	2.79	3.88	.03	47.14
	>80	4,156	1.30	1.84	.01	23.52
	Total	9,836	2.55	4.13	.01	69.67
EUR	1-20	3,709	11.65	10.39	.15	86.22
	21-80	7,775	7.42	7.12	.04	51.83
	>80	4,,113	4.15	4.04	.01	32.46
	Total	15,597	7.56	7.89	.01	86.22
MEA	1-20	234	5.78	6.29	.13	40.63
	21-80	521	3.06	3.77	.04	35.10
	>80	706	2.16	2.32	.01	16.12
	Total	1,461	3.06	3.95	.01	40.63

Source: Compiled by the authors based on data from Booking.com

The results by managerial form and hotel size, show the same pattern, the smaller the hotel is, the higher RpR have, as can be seen in Table 7 with the ANOVA test, showing all the results statistically significant ($p < .001$) in all cases.

Table 7. Descriptive statistics of RpR by independent and chain hotels and hotel size

		N	Mean	SD	Min	Max
Independent	1-20	6,853	8.95	9.40	.13	86.22
	21-80	13,330	5.32	6.31	.03	51.83

	>80	7,139	2.40	3.27	.01	32.46
	Total	27,322	5.47	7.06	.01	86.22
	1-20	123	9.10	9.16	.25	40.89
Chain	21-80	1,687	5.47	5.32	.04	40.67
	>80	4,763	2.64	2.99	.01	26.79
	Total	6,573	3.49	4.18	.01	40.89

Source: Compiled by the authors based on data from Booking.com

4.4. RpR by countries

By countries we can see in Table 7 and in Figure 1 that the highest RpR mean is in European countries, the first non-european country is Costa Rica, in the 14th position. Both, Table 7 and Figure 1, only include countries with more than 100 hotels in our sample (45) although the global sample includes 66 countries.

Table 7. Descriptive statistics of RpR by countries

Country	N	Mean	SD	Min	Max
Netherland	566	14.12	11.73	0.15	86.22
Italy	2359	9.04	9.74	0.02	83.50
United Kingdom	1152	8.84	7.93	0.09	47.15
Spain	1130	8.72	8.08	0.02	59.78
France	1929	8.35	6.61	0.11	51.92
Poland	441	7.51	6.42	0.24	51.83
Germany	1480	7.37	6.88	0.05	59.00
Ireland	300	7.27	6.69	0.28	34.19
Portugal	490	6.74	6.98	0.06	47.94
Switzerland	484	6.58	6.23	0.06	33.71
Greece	537	6.57	7.03	0.01	38.90
Russia	1486	6.03	7.50	0.02	64.00
Austria	822	5.58	7.12	0.04	52.29
Costa Rica	128	5.42	5.47	0.16	28.38
Turkey	1520	5.35	6.47	0.02	54.60
Chile	277	5.21	6.27	0.06	56.10
Argentina	579	5.18	5.39	0.07	35.10
New Zealand	195	5.17	4.90	0.40	30.80
Romania	399	5.08	5.23	0.09	33.50
Denmark	285	4.82	5.42	0.05	37.71
Norway	217	4.61	4.90	0.32	26.48
Australia	437	4.54	4.33	0.15	26.69
UAE	311	4.30	2.72	0.44	16.12
Canada	500	4.07	4.27	0.04	31.92
Cambodia	368	3.98	5.36	0.02	35.00
Brazil	892	3.95	4.38	0.06	64.75

Colombia	820	3.94	4.80	0.05	47.83
Taiwan	831	3.94	5.92	0.04	63.08
Israel	320	3.78	5.84	0.01	40.63
Singapore	222	3.46	3.39	0.15	26.91
Thailand	1478	3.26	4.74	0.02	69.67
Peru	480	3.21	5.12	0.02	58.95
South Africa	348	3.13	3.74	0.08	18.72
Mexico	650	3.03	4.55	0.02	43.88
Morocco	154	2.81	3.38	0.05	25.00
Vietnam	1327	2.72	4.33	0.01	29.72
United States	2180	2.56	4.09	0.01	78.46
Malaysia	653	2.46	3.58	0.06	40.70
Japan	762	2.15	3.86	0.02	38.45
Nepal	104	1.94	2.59	0.05	12.22
Indonesia	748	1.93	3.14	0.03	35.65
India	958	1.58	2.84	0.02	27.78
South Corea	241	1.40	1.85	0.04	14.39
Egypt	166	1.38	2.14	0.04	21.52
China	1486	0.88	2.06	0.01	35.76

Source: Compiled by the authors based on data from Booking.com

Insert Figure 2

5. Conclusions

The results show that Booking.com has a dominant position in Europe, with relevant countries at the international tourist sphere, such as Netherlands, Italy, United Kingdom, Spain, and France with the highest RpR mean worldwide..

The expansion of Booking.com Worldwide has been uneven in non-european countries, with the highest implementation of Booking.com in Costa Rica, Chile, and Argentine (AME) followed by New Zealand and Australia (ASP). Other relevant markets but with a small share are United States, Mexico, Brazil, China, India or Vietnam, among others. The great presence of Expedia brands in AME, Priceline in United States of America (Blomberg-Nygard and Anderson, 2016) and websites like Agoda and Ctrip which both have a high presence of hotels in ASP, specially in China (Zhou et al., 2014) or HotelTravel in India (Blomberg-Nygard and Anderson, 2016) can explain these differences.

The low levels of RpR in most of the countries outside of Europe should not be considered as a failure of Booking.com's expansion strategy. Booking.com has achieved a market share of 62% among OTAS with an RpR of around 8 in the most European countries (Schegg, 2016). Therefore, levels of 4 in the RpR of Australia, Canada or Brazil, could equal market shares of around 30% in those markets. In

others such as Thailand, México or United States where the RPR is close to 3, Booking.com could have market shares of 20-25%.

The results show how there are factors that clearly influence the use level of Booking.com, not only in Europe, where the company is strongly positioned for more than 10 years, the results are repeated in geographical areas where the implementation of Booking.com is relatively recent. The three identified factors deserve a deeper discussion:

a) Independent hotels Vs Chain hotels: It is logical to expect that hotels belonging to hotel chains (especially in the case of medium and large hotel chains) have the capacity to make important investments to promote direct sales, either through the hotel's website, customer loyalty programs or other alternative channels. Numerous studies in the academic literature identified by (Ben Aissa and Goaied, 2016) highlight that those hotels belonging to hotel chains are much more profitable than the independent ones. Probably, the profitability improvement comes from avoiding paying commissions to agencies like Booking.com. Moreover, small hotels usually do not have a large reservation engine and use the Booking.com engine (Balagué et al., 2016) pasting the code link into their websites (Booking, 2017b) so customers book through this OTAs or through their websites but with the OTA engine which count also as a reservation done through Booking.com.

b) Small hotels Vs Large hotels: A small hotel usually implies limited financial resources and great difficulty to have a large and experienced sales and marketing team. In these circumstances, the easiest and most practical strategy is to use intermediaries such as Booking.com to channel hotel sales, assuming minimum commissions to be paid in this websites or even paying additional commissions to appear in the top of rankings. It could be related also to the findings of (Ben Aissa and Goaied, 2016) that confirm that large hotels get high occupancy and better sales revenues.

c) Low category Vs High category: The results confirm that hotels with highest categories sell fewer rooms through Booking.com. Otherwise, hotels with lowest categories have bigger dependency on Booking.com because they have a higher RpR. This relation could be explained by the limitations in financial and staff resources commented previously for the case of small and large hotels. However there is no mean difference for 1-star and 2-stars hotels, which could be due to the fact that the features of those hotels are very similar (Martin-Fuentes et al., 2018).

The concerns about possible bias, because of reviewers could be much more likely to post reviews if they are extremely satisfied or disappointed, have been minimized after analyze a previous 662,991 reviews database and search for a relationship between scores and RpR, which we did not detect. Even if extremely disappointed customers trend to leave more reviews, there is such a small percentage of very negative reviews that the global effect would be not significant. In the case of extremely satisfied customers, if there was a significant bias, we would have detected that hotels with better scores tend to have a higher RpR, which has not happened either.

Differences observed shows that dependency on Booking.com varies within the same tourist destination. This study draws a situation in which hotels with a smaller structure are forced (or opt voluntarily) to use Booking.com with greater intensity. What hotels should consider is whether it really is worth making that economic effort to increase direct sales or if it is preferable to continue paying commissions close to 20% to intermediaries such as Booking.com. Nonetheless, increase the percentage of direct sales does not depend exclusively on the economic investment made to enhance this channel. Many times the use of appropriate strategies can improve these percentages without involving direct economic investments.

6. Limitations and future research

This study presents several limitations, since we worked on estimates and also focus on a specific profile of hotels in big destinations. Moreover, we used the number of reviews as a proxy to measure actual hotel sales and assumed homogeneous occupancy levels and average stay. Although this can be considered as a valid approach, it may cause an information distortion, particularly when we do not have information about response rates, occupancy levels or average stay length.

Intercultural differences between countries could imply differences in Booking.com surveys response rates, which could explain partially differences in RpR by country. However, the variations detected in the RpR according to size, hotel chain and stars are almost identical in all the defined geographical areas and in practically all of the countries analyzed. Therefore, the robustness and uniformity of these data makes us think that these factors are really significant in the level of use of Booking.com by hotels.

This study is only a first approximation to how the RpR can be used to determine the level of sales that an OTA has in a single hotel or in a group of hotels. It would be advisable to carry out a parallel study, obtaining the RpR of the most important OTAs in each geographical area to obtain a more complete view of market share by each OTA. It could apply to other OTAs that have a similar reviews system, like Expedia, Agoda or HRS. Moreover, the RpR figures can be tracked over time to determine changes in the hotel sales mix. Going further, automating and monitoring the entire data extraction process, public authorities or private entities could know in real time the weight of each OTA in the sales of hotels around the world.

It would also be very interesting to carry out an in-depth study on the response rate of the Booking.com surveys, which would allow to know more accurately the relationship between the number of reviews and sales. This could be developed through surveys that ask about the number of reservations made with Booking.com, total reservations, occupancy and length of stay. However, it is very difficult to obtain this information from hotels, which are usually reluctant to share this information.

7. References

Abrate, G., & Viglia, G. (2016). Strategic and tactical price decisions in hotel revenue management. *Tourism Management*, *55*, 123–132. https://doi.org/10.1016/j.tourman.2016.02.006

Balagué, C., Martin-Fuentes, E., & Gómez, M. J. (2016). Reliability of authenticated versus nonauthenticated hotel reviews: Tripadvisor and booking. Com case. *Cuadernos de Turismo*, (38).

Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, *53*, 125–131. https://doi.org/10.1016/j.tourman.2015.09.020

Ben Aissa, S., & Goaied, M. (2016). Determinants of Tunisian hotel profitability: The role of managerial efficiency. *Tourism Management*, *52*, 478–487. https://doi.org/10.1016/j.tourman.2015.07.015

Bennett, M. M., & Lai, C.-W. K. (2005). The impact of the internet on travel agencies in Taiwan. *Tourism and Hospitality Research*, *6*(1), 8–23.

Berne, C., Garcia-Gonzalez, M., & Mugica, J. (2012). How ICT shifts the power balance of tourism distribution channels. *Tourism Management*, *33*(1), 205–214. https://doi.org/10.1016/j.tourman.2011.02.004

Bjørkelund, E., Burnett, T. H., & Nørv\a ag, K. (2012). A study of opinion mining and visualization of hotel reviews. In *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services* (pp. 229–238). ACM.

Blomberg-Nygard, A., & Anderson, C. K. (2016). United Nations World Tourism Organization study on online guest reviews and hotel classification systems: An integrated approach. *Service Science*, *8*(2), 139–151. https://doi.org/10.1287/s erv.2016.0139

Booking. (2013). Booking.com Launches 'Booking.yeah', Its First-Ever Brand Campaign, Created for the U.S. market.

Booking. (2017a). Reviews: How does it work?

Booking. (2017b). The Booking Button. Retrieved 27 December 2017, from https://www.booking.com/booking_button.html

Borges, I. R., Pereira, G. M., Matos, C. A. de, & Borchardt, M. (2015). Analysis of the relationship between the satisfaction of consumers and the prices offered on site booking.com. *Tourism & Management Studies*, *11*(2), 64–70.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management*, *29*(4), 609–623. https://doi.org/10.1016/j.tourman.2008.01.005

Cantallops, A. S., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41–51.

Carroll, B., & Siguaw, J. (2003). The evolution of electronic distribution: Effects on hotels and intermediaries. *The Cornell Hotel and Restaurant Administration Quarterly*, *44*(4), 38–50. https://doi.org/10.1016/S0010-8804(03)90257-6

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales : Online Book Reviews, *XLIII*(August), 345–354.

Chiappa, G. D., & Dall'Aglio, S. (2012). Factors influencing Travellers' e-Ratings and e-Complaints about Hotel Services: Insights from an Italian Tourism Destination. In

Information and Communication Technologies in Tourism 2012 (pp. 448–459). Springer, Vienna. https://doi.org/10.1007/978-3-7091-1142-0_39

Díaz, M. R., & Rodríguez, T. F. E. (2017). Determining the reliability and validity of online reputation databases for lodging: Booking.com, TripAdvisor, and HolidayCheck. *Journal of Vacation Marketing*, 1356766717706103. https://doi.org/10.1177/1356766717706103

European Commission. (2017). Report on the monitoring exercise carried out in the online hotel booking sector by EU competition authorities in 2016. Retrieved from http://ec.europa.eu/competition/ecn/hotel_monitoring_report_en.pdf

Ford, R. C., Wang, Y., & Vestal, A. (2012). Power asymmetries in tourism distribution networks. *Annals of Tourism Research*, *39*(2), 755–779. https://doi.org/10.1016/j.annals.2011.10.001

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291–313.

Garrigos-Simon, F. J., Galdon, J. L., & Sanz-Blas, S. (2017). Effects of crowdvoting on hotels: the Booking.com case. *International Journal of Contemporary Hospitality Management*, *29*(1), 419–437. https://doi.org/10.1108/IJCHM-08-2015-0435

Ghose, A., & Ipeirotis, P. G. (2006). Designing ranking systems for consumer reviews: The impact of review subjectivity on product sales and review quality. In *Proceedings of the 16th annual workshop on information technology and systems* (pp. 303–310).

Govers, R., & Go, F. M. (2004). Projected destination image online: Website content analysis of pictures and text. *Information Technology & Tourism*, *7*(2), 73–89. https://doi.org/10.3727/1098305054517327

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*(1), 38–52.

Ip, C., Leung, R., & Law, R. (2011). Progress and development of information and communication technologies in hospitality. *International Journal of Contemporary Hospitality Management*, *23*(4), 533–551.

Kwok, L., Xie, K. L., & Richards, T. (2017). Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals. *International Journal of Contemporary Hospitality Management*, *29*(1), 307–354.

Law, R., Leung, R., Lo, A., Leung, D., & Fong, L. H. N. (2015). Distribution channel in hospitality and tourism: Revisiting disintermediation from the perspectives of hotels and travel agencies. *International Journal of Contemporary Hospitality Management*, *27*(3), 431–452. https://doi.org/10.1108/IJCHM-11-2013-0498

Levy, A. (2013). Booking.com Challenging Parent Priceline in U.S. Travel. Retrieved from https://www.bloomberg.com/news/articles/2013-01-22/booking-com-challenging-parent-priceline-in-u-s-travel

Lu, Q., Xiao, L., & Ye, Q. (2012). Investigating the impact of online word-of-mouth on hotel sales with panel data. In *Management Science and Engineering (ICMSE), 2012 International Conference on* (pp. 3–9). IEEE.

Lu, Q., Ye, Q., & Law, R. (2014). Moderating effects of product heterogeneity between online word-of-mouth and hotel sales. *Journal of Electronic Commerce Research*, *15*(1), 1–12.

Mariani, M. M., & Borghi, M. (2018). Effects of the Booking.com rating system: Bringing hotel class into the picture. *Tourism Management*, *66*, 47–52. https://doi.org/10.1016/j.tourman.2017.11.006

Martin-Fuentes, E. (2016). Are guests of the same opinion as the hotel star-rate classification system? *Journal of Hospitality and Tourism Management*, *29*, 126–134. https://doi.org/10.1016/j.jhtm.2016.06.006

Martin-Fuentes, E., Fernandez, C., Mateu, C., & Marine-Roig, E. (2018). Modelling a grading scheme for peer-to-peer accommodation: Stars for Airbnb. *International Journal of Hospitality Management*, *69*. https://doi.org/10.1016/j.ijhm.2017.10.016

Mellinas, J. P. (2017). Dependency of Spanish urban hotels on Booking.com. In *9th World Conference for Graduate Research in Tourism Hospitality and Leisure* (pp. 177–182). Cartagena, Spain: Anatolia.

Mellinas, J. P., Martínez María-Dolores, S.-M., & Bernal García, J. J. (2015). Booking.com: The unexpected scoring system. *Tourism Management*, *49*, 72–74. https://doi.org/10.1016/j.tourman.2014.08.019

Mellinas, J. P., Martínez María-Dolores, S.-M., & Bernal García, J. J. (2016). Effects of the Booking.com scoring system. *Tourism Management*, *57*, 80–83. https://doi.org/10.1016/j.tourman.2016.05.015

Murphy, C. (2017). Report: 78% of All Online Hotel Reviews Come From the Top Four Sites. Retrieved from https://learn.revinate.com/blog/report-78-of-all-online-hotel-reviews-come-from-the-top-four-sites

Öğüt, H., & Onur Taş, B. K. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *The Service Industries Journal*, *32*(2), 197–214.

Pal, S. K., & Mishra, P. (2017). Portfolio of Online Distribution Channels across Mid-Market Hotels: An Evaluative Review. *ENLIGHTENING TOURISM. A PATHMAKING JOURNAL*, *7*(1), 19–35.

Pearce, D. G. (2008). Beyond Tiers: A Network Approach to Tourism Distribution. *Tourism Analysis*, *13*(5–1), 517–530. https://doi.org/10.3727/108354208788160496

Priceline. (2017). Retrieved 20 October 2017, from http://www.pricelinegroup.com/booking-com/.

Radojevic, T., Stanisic, N., & Stanic, N. (2015a). Solo travellers assign higher ratings than families: Examining customer satisfaction by demographic group. *Tourism Management Perspectives*, *16*, 247–258. https://doi.org/10.1016/j.tmp.2015.08.004

Radojevic, T., Stanisic, N., & Stanic, N. (2015b). Solo travellers assign higher ratings than families: Examining customer satisfaction by demographic group. *Tourism Management Perspectives*, *16*, 247–258. https://doi.org/10.1016/j.tmp.2015.08.004

Radojevic, T., Stanisic, N., & Stanic, N. (2017). Inside the Rating Scores: A Multilevel Analysis of the Factors Influencing Customer Satisfaction in the Hotel Industry. *Cornell Hospitality Quarterly*, *58*(2), 134–164. https://doi.org/10.1177/1938965516686114

Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In *The Economics of the Internet and E-commerce* (pp. 127–157). Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/doi/pdf/10.1016/S0278-0984(02)11030-3

Richard, B., Perry, W. P., & Ford, R. C. (2016). Crowdsourcing in the Lodging Industry: Innovation on a Budget. In *Open Tourism* (pp. 79–94). Springer, Berlin, Heidelberg.

Schegg, R. (2016). European Hotel Distribution Study. Results for the Reference Year 2015. Retrieved 19 July 2017, from https://www.oehv.at/Lobbying/Tourismusmarkt-Osterreich/Hotrec-Umfrage-Hoteldistribution/final_public_schegg_2016_european_hotel_distributi.aspx

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, *32*(5), 608–621.

Stanisic, N. (2016). *Recent Trends in Quantitative Research in the Field of Tourism and Hospitality* (SSRN Scholarly Paper No. ID 2875849). Rochester, NY: Social Science Research Network.

Talluri, K. T., & Van Ryzin, G. J. (2006). *The theory and practice of revenue management* (Vol. 68). Springer Science & Business Media.

Toh, R. S., Raven, P., & DeKay, F. (2011). Selling rooms: Hotels vs. third-party websites. *Cornell Hospitality Quarterly*, *52*(2), 181–189.

Tse, A. C. (2003). Disintermediation of travel agents in the hotel industry. *International Journal of Hospitality Management*, *22*(4), 453–460. https://doi.org/10.1016/S0278-4319(03)00049-5

Viglia, G., Furlan, R., & Ladron-de-Guevara, A. (2014). Please, talk about it! When hotel popularity boosts preferences. *International Journal of Hospitality Management*, *42*, 155–164.

Wetzer, I. M., Zeelenberg, M., & Pieters, R. (2007). "Never eat in that restaurant, I did!": Exploring why people engage in negative word-of-mouth communication. *Psychology & Marketing*, *24*(8), 661–680.

Whitty, M. T., & Joinson, A. (2008). *Truth, Lies and Trust on the Internet* (1 edition). London ; New York: Routledge.

Ye, Q., Gu, B., Chen, W., & Law, R. (2008). Measuring the value of managerial responses to online reviews-A natural experiment of two online travel agencies. *ICIS 2008 Proceedings*, 115.

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, *28*(1), 180–182. https://doi.org/10.1016/j.ijhm.2008.06.011

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, *27*(2), 634–639. https://doi.org/10.1016/j.chb.2010.04.014

Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, *36*(3), 6527–6535. https://doi.org/10.1016/j.eswa.2008.07.035

Yoo, K. H., & Gretzel, U. (2008). The influence of involvement on use and impact of online travel reviews. In *Hospitality Information Technology Association (HITA) conference* (pp. 15–16).

Zhou, L., Ye, S., Pearce, P. L., & Wu, M.-Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality Management*, *38*, 1–10. https://doi.org/10.1016/j.ijhm.2013.12.004

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148.