It’s not all about price. The determinants of occupancy rates in peer-to-peer accommodation: A methodological contribution

Published Version:
It’s not all about price. The determinants of occupancy rates in peer-to-peer accommodation: A methodological contribution / Veronica Leoni, Paolo Figini, William Nilssom. - In: INTERNATIONAL JOURNAL OF CONTEMPORARY HOSPITALITY MANAGEMENT. - ISSN 0959-6119. - STAMPA. - 32:4(2020), pp. 1693-1711. [10.1108/IJCHM-05-2019-0464]

Availability:
This version is available at: https://hdl.handle.net/11585/774574 since: 2021-03-01

Published:
DOI: http://doi.org/10.1108/IJCHM-05-2019-0464

It's not all about price. The determinants of occupancy rates in p2p accommodation: a methodological contribution.

Structured Abstract

Purpose

The paper identifies the key drivers of occupancy rates in peer-to-peer accommodation.

Design/methodology/approach

The applied methodology fits the specific characteristics of this market segment: (i) the peculiar distribution of the occupancy rate (a ratio characterised by a large share of zeros) requires the adoption of a mixed discrete-continuous model; (ii) the insidious issue of price endogeneity is dealt with a control function approach; (iii) the econometric specification takes into account the monopolistic competition, the relevant market regime in the hospitality industry. The model is tested on Airbnb listings in the Balearic Islands (Spain).

Findings

The occupancy rate of peer to peer properties in the Balearic Islands strongly depends on their geographical location and online reputation. There is a qualitative difference between two groups: listings with positive occupancy rates, which demand tends to be inelastic, and listings with zero occupancy. We found that the price is a not a statistically significant determinant of the group membership.

Originality/value

This paper applies a zero-inflated beta model, never used in previous analyses of occupancy rates, to provide a benchmark for future studies. This procedure allows the estimation of unbiased marginal effects. It thus offers important technical and managerial implications, as a wrong understanding of how occupancy depends on price would deliver ineffective managerial decisions. This paper highlights the importance of methodological choices, since coefficients are highly sensitive to misspecifications of the model.

Introduction and Positioning

A growing interest in the 'Sharing Economy' (SE), also referred as peer-to-peer (p2p) economy, has been documented in academic research, media and public discourse. In this respect, the tourism and travel sector is at the forefront, with relevant experiences such as Blabla Car in road transport and Airbnb in accommodation services.

Most of the recent literature focuses on Airbnb, the posterchild of p2p economy and recently valued at US$38 billion[1]. The relevance of Airbnb and its presumed disruptive effect on the hospitality market have fomented a widespread academic interest and triggered investigation, among other things, into the motivations behind people's participation in sharing activities, the impact on traditional service suppliers, and the analysis of revenue performance, price dynamics and occupancy rates (OR).

This paper contributes to this last topic: the determinants of demand in the SE. First, the paper tackles some of the data and methodological issues at stake when working with Airbnb (and similar marketplaces). Airbnb data have three main peculiarities: (i) most of the listings have an on/off nature: the listing can be 'off' (unavailable) if the host has 'blocked' it (using Airbnb terminology) on purpose in order to avoid bookings; (ii) most of the listings are composed of only one unit (unlike multi-room hotels) and hence, in a single day, the OR can only be 0 or 1. Therefore, ORs can only be computed over a meaningful aggregate (either over time or geographically); (iii) because of its fractional nature and the massive presence of zeros (that is, listings that are never booked during the period under investigation), the distribution of OR is peculiar, thus generating critical concerns when applying basic linear or pure binary models. To tackle these features, for which it would be

[1] Forbes: https://www.forbes.com/sites/greatspeculations/2018/05/11/as-a-rare-profitable-unicorn-airbnb-appears-to-be-worth-at-least-38-billion/

erroneous to adopt the same approach used to study OR in traditional accommodation, we apply a mixed discrete-continuous model.

The OR is a key indicator of performance in the hospitality industry, together with the average daily rate (ADR) and the revenue per available room (REVPAR). While it is likely that p2p hosts aim at maximizing REVPAR more than occupancy, OR measures the quantity sold and it is hence a direct indicator of demand, which can be explained by price, among other things. Since the focus of this study is on the determinants of demand in the SE (not on performance), we will proxy it with the OR.

When investigating demand, price endogeneity is one of the critical estimation issues to address and a well-discussed topic in tourism economics and other fields. Many studies have shown that not accounting for endogeneity leads to biased estimates of price elasticity (Kim and Uysal, 1997; Pekgün *et al*., 2013; Mumbower *et al*., 2014; Lurkin *et al*., 2017). Assuming that the price is exogenous could engender inaccurate conclusions and policy implications. To the best of our knowledge, none of the existing studies on OR in the hospitality sector correctly treats endogeneity. Only Li *et al.* (2015) tested the robustness of their results against potential endogeneity concerns. Hence, the second novel contribution of this work is the proper treatment of endogeneity through a control function approach, applied for the first time in the hospitality sector. The implementation of this approach was possible because the characteristics and the precise geographical location of listings were used, together with daily data on price and OR.

The third novel contribution is to provide a unifying theoretical framework to guide the empirical analysis. Unlike most previous studies, this paper builds the econometric specification upon the model of monopolistic competition, the market regime where Airbnb listings compete (Gunter and Önder, 2018; Zekan *et al*., 2019). Hence, the hypotheses are directly derived from the theory, and the empirical analysis assesses its validity.

The paper analyses Airbnb listings in the Balearic Islands, one of the most important tourism destinations worldwide. By delivering accurate and unbiased estimations of ORs through the adopted methodological approach, this paper provides insights into the real extent of Airbnb demand and its relationship with price. Hence, the study's final contribution is also empirical, adding to the scientific debate on the relevance of SE in such a mature and international destination.

Due to space constraints, the literature review related to the OR in the hospitality industry is summarised in Table 1, which shows the case studies, the methodologies and the main findings. As regards the determinants of OR in the Sharing Economy, which is the focus of our study, the main contributions of the related literature are here recalled. The emerging SE business model calls in fact for specific and novel analysis of ORs in p2p marketplaces. Van der Borg *et al.* (2017) studied the impact of community aspects (e.g. the '*superhost'* label) and of the characteristics of Airbnb listings on OR in the Veneto region, Italy. Li *et al*. (2015) highlighted the difference between non-professional and professional hosts, finding that properties managed by professional hosts achieve on average 15.5% higher ORs than properties managed by non-professional hosts. Similarly, Gunter and Önder (2018) analysed Airbnb in Vienna through a cluster robust OLS model to identify the determinants of ORs. They found that demand for Airbnb is price-inelastic, thus leading to larger revenues via price hiking. They also found a positive relationship between ORs and host responsiveness, the number of photos, the listing size and proximity to the city centre. This paper partially builds upon Gunter and Önder (2018) by refining the methodology used and applying it to the case of Airbnb in Balearics.

[Insert Table 1 about here]

Theoretical framework and research hypotheses

From a theoretical perspective, the monopolistic competition is the relevant market structure for accommodation in private houses because of the strong differentiation of listings in markets where many suppliers compete (Gunter and Önder, 2018; Zekan *et al*., 2019).

[Equation 1] $$Q = S\,[\frac{1}{n} - b(p - P)]$$

Consistent with the basic model of monopolistic competition (Salop, 1979) recalled in equation [1], this paper assumes that the quantity (Q) of a listing demanded by tourists negatively depends on its own price (p), positively on the price (P) set by the 'competitors', i.e., other listings with similar characteristics and location; negatively on the number of competitors (n). The parameter "b" is the slope of the demand function. As individual supply is fixed by room capacity, and assuming that the market size (S) is also fixed in the (short) period under consideration, it is easy to consider demand in relative terms (Q/S) and proxy it with OR. Accordingly, equation [1] constitutes the theoretical model to be estimated, and the following research hypotheses can be tested:

Hypothesis 1 (H1). Price has a negative effect on the occupancy rate.

The law of demand implies that the price is inversely related to the quantity demanded. However, the price should also work as a quality signal and have a positive effect. Since it is possible to disentangle and control for quality (discussed in hypotheses 4 and 5), H1 expects that the price has a negative effect on the OR.

However, the importance of price has often been overestimated (Lockyer, 2005) as previous research (Van der Borg, 2017; Gunter and Önder, 2018) highlighted that Airbnb demand can be quite price inelastic. Hence, the precise estimation of how OR is sensitive to price is key to understanding the tourist property's performance. Their purchase intentions depend on the perceived gain (or loss) with respect to the reference price (the price they have stored in mind as the 'correct' one, Oh, 2003).

People tend to use Airbnb only when it involves a real cost saving (Tussyadiah and Pesonen, 2016) and its value perception is higher for very price-sensitive users (Liang *et al*. 2018).

Hypothesis 2 (H2): The average price of (local) competitors has a positive effect on the individual occupancy rate.

A listing in a monopolistically competitive industry is expected to sell more the higher the price set by its rivals, which is tantamount to assuming that given a listing's own price, individual market share increases with the price of (local) competitors. This also relates to the idea recalled in H1, since the price of competitors might work as the reference price (Viglia *et al*., 2016).

Hypothesis 3 (H3): The number of listings available in the same area is negatively linked to the occupancy rate.

This hypothesis stems directly from the Salop model (1979). The number of available listings in the same area negatively affects individual demand since, at least in the short run, the market size is fixed. This means that hosts can gain market share only at the expense of other listings: stronger levels of competition negatively affect the OR.

While these three assumptions directly stem from the issue of variety and horizontal differentiation of the Salop model, the next three assumptions address the role played by quality:

Hypothesis 4 (H4). The volume and the rating scores of reviews are positively associated with the occupancy rate.

After controlling for the price, it is likely that listings with better reputation have higher demand. The impact of online reviews on sales is a hot topic in tourism economics and management literature (Ye *et al.,* 2009; Öğüt and Tas, 2012; Mauri and Minazzi, 2013 among the many). Both the number of reviews and the average rating score have significant effects on hotels' ORs (Viglia *et al.,* 2016) and

Airbnb listings' ORs (Van der Borg *et al.*, 2017; Gunter and Önder, 2018). This hypothesis is consistent with such literature.

Given the high degree of human interaction in p2p markets and the perception of higher risk compared to traditional accommodation, personal reputation is paramount in building listings' popularity, outweighing the importance of product description (Mauri *et al*., 2018). Consequently, we expect that variables associated with human interaction between hosts and guests (e.g., the number of photographs of the property, the label of 'super host' provided by Airbnb, the responsiveness of the host to booking requests) are all associated with higher ORs.

Hypothesis 5 (H5). The distance to the beach has a negative impact on the occupancy rate.

While in the Salop model (1979) product differentiation is horizontal, goods and services are also vertically differentiated. Quality in accommodation has two dimensions. The first one is already discussed in H4; the second concerns the location (in terms of distance/proximity to the main tourism attraction), with non-trivial implications for customer purchase intention and on the evaluation of the stay experience (Tussyadiah, 2016; Yang *et al.*, 2017). In this study, being the Balearics predominantly a seaside destination, it is reasonable to expect that the further the listing from the beach, *ceteris paribus*, the lower the OR.

Hypothesis 6 (H6) The level of local performance is positively linked to the occupancy rate.

According to Porter (1998), synergies within a production cluster produce external economies of scale and enhance the productivity and innovation of local firms, as well as their economic performance. Peiró-Signes *et al.*, (2015) found that hotels within a geographical cluster economically outperform those outside it. Since in tourism geographical concentration often overlaps with geographical proximity to the main attraction, we attempt to disentangle these two effects.

While in H5 we test the effect of being close to the main attraction, and in H3 we study the effect of local competition on the OR, H6 aims at discovering the effect of being in a good location (in terms of ‘highly performing district’). We expect that listings located in high-performing areas, *ceteris paribus*, will have higher ORs.

Data and Methodology

Data

This paper investigates the OR determinants for Airbnb listings in the Balearic Islands, using data for July and August 2016. The relevance of this study stems from jointly considering the world leader in SE markets and one of the most popular international destinations in the 'sea and sun' segment. The Balearic Islands, located in the Mediterranean Sea, recorded more than 20 million arrivals in 2018. Tourists travel to the Balearics mainly for leisure purposes (89.7%), and the tourism destination is characterised by a strong seasonality (arrivals are concentrated in the period May–September, AETIB, 2018)[2]. Tourism generates almost 45% of the Balearic GDP and 3.2 of every 10 jobs (IMPACTUR, 2014)[3].

The dataset includes approximately 20,000 Airbnb listings[4], of which we considered only those active at least in June, July and August 2016 (15,203). Descriptive statistics for the variables included in the analysis are reported in Table 2.

[Insert Table 2 about here]

We generate monthly performance metrics using daily information on price and booking status (i.e., day booked, day available or day blocked[5]). The dataset also includes information about the listing's features, the host's characteristics, the geographical location and the most important indicators of online reputation (the number of reviews, the overall rating score and the host attribute as 'super-host'[6] (Teubner *et al.*, 2016; Liang *et al.*, 2017)). As ratings show low variance and the typical J-

[2] AETIB webpage: El turisme a les Illes Balears. Anuari 2018 http://www.caib.es/sites/estadistiquesdelturisme/es/inicio-23165/?campa=yes

[3] IMPACTUR: https://www.exceltur.org/wp-content/uploads/2015/10/IMPACTUR-Baleares-2014-informe-completo.pdf

[4] The dataset has been supplied by Airdna (https://www.airdna.co/).

[5] According to the Airbnb definition, blocked days are days when the host deliberately decides not to accept any booking.

[6] '*Superhost'* description: https://www.airbnb.co.uk/help/article/828/what-is-a-superhost

shape found in online reviews (Hu *et al.*, 2009), a set of dummy variables was generated: ‘high_rate’, ‘medium_rate’ and ‘low_rate’ (description available in Table 2). As per Airbnb policy, listings with less than three reviews appear as ‘never_rated’.

The listings’ longevity expresses the months of presence on the platform since the date of registration. We also control for hosts with more than one property (dummy variable) and the number of properties managed (discrete variable). In addition, we consider the hosts’ managerial decisions, including the minimum stay required per booking (discrete variable) and whether the booking request is automatically accepted (dummy variable).

Methodology

As recalled in the introduction, the study of the OR determinants must tackle a few estimation issues stemming from the fractional nature of the dependent variable, from endogeneity between OR and price and from the measurement of geographical location. In this section, we carefully discuss these problems.

Occupancy rate

The OR is the ratio between the number of days booked and the number of days the listing was available (that is, the number of days booked plus the number of days unsold); hence, it has a fractional nature and the denominator does not include ‘blocked days’. The interpretation of blocked days is non-trivial, since it could signal that either the service was available and sold on other platforms/channels or that the owner (or his/her acquaintances) used the apartment, or simply that it was undertaking preparation for future guests.

[Insert Figure 1 about here]

As Figure 1 suggests, approximately 34% of observations have OR=0 (listings never booked[7]). Its peculiar distribution makes the use of linear regressions unattractive (Buis, 2012); as the literature suggests, a zero-one-inflated beta model (from now on ZOIB) is the best option (Ferrari and Cribari-Neto, 2004; Ospina and Ferrari, 2012). ZOIB works as a mixed continuous-discrete model involving three steps. First, it structurally distinguishes between zeros and non-zeros. Second, it distinguishes between ones and non-ones. Third, it analyses listings with an OR in the (0; 1) range. In this paper, the second step (one-inflation) is omitted because the limited number of observations with OR=1 (2.7% of the sample, see Figure 1) makes this step unnecessary.

The Beta distribution describes the quantitative component of the model (third step), following equation [2].

[Equation 2] $$P(a,b) = \frac{\Gamma[a+b]}{\Gamma(a)\Gamma(b)}\, y^{a-1}\, [1-y]^{b-1}$$ [8]

The binary part (first step) is instead treated with a logistic regression, following equation [3].

[Equation 3] $$y = \frac{exp\frac{x-a}{b}}{1+exp\frac{x-a}{b}}$$

These models are run simultaneously and have their own set of predictors and coefficients.

Price

The original dataset was a panel of daily prices and their associated status (i.e., booked, non-booked/available, or blocked). The price is aggregated over time and the panel transformed into a

[7] The histogram shows the aggregated value of the OR in July and August 2016.

[8] α and β are two positive parameters controlling the shape of the distribution.

cross-section: the average price (in natural logarithms) over the period under investigation, regardless of the booking status, is the variable used to study the effect of price on the OR. The main disadvantage of this aggregation is the loss of information, as we cannot capture intra-listing price variations or the effect of micro-seasonality or special events.

We did not use the ADR because it would return a missing value in the case of never-booked properties: neglecting this part of the sample would lead to an incomplete and erroneous understanding of the phenomenon. In addition, ADR represents the price at which the room is 'rented' and entails a severe endogeneity problem. The average listed price, on the contrary, is the price at which the host is willing to rent the property online. The endogenous nature of price is a well discussed topic and many studies have shown that price coefficients are biased if endogeneity is not accounted for (Kim and Uysal, 1997; Pekgün *et al.,* 2013; Mumbower *et al*., 2014; Lurkin *et al.,* 2017). Assuming that the price is exogenous could lead to inaccurate conclusions and policy implications. This is particularly relevant in the accommodation industry, which is characterised by a capacity constraint. In the short run, the only way to react to demand shocks is through price changes, since it would take time for the supply to adjust to demand through an increase in hotel capacity. Hence, not only has the price a strategic dimension linked to the static characteristics of the service, but it also entails a tactical dimension as it adjusts depending on demand changes (Abrate and Viglia, 2016).

The canonical methods proposed for fractional models with continuous Endogenous Explanatory Variables (EEV; Wooldridge, 2011) are inconvenient for p2p accommodation as they do not consider the zero-inflation characteristics of such a market. Therefore, we opted for the control function approach, which relies on the same identification conditions (Wooldridge, 2015) and provides a better approximation of the marginal effects than other conventional methods (Wooldridge, 2011). Finding valid instruments to treat endogeneity is always complicated. A valid instrument requires two binding conditions: (i) be statistically correlated with the EEV; and (ii) be exogenous to all the other unobserved factors (French and Popovici, 2011). We choose three instruments: the daily, weekly and monthly prices set by the host when publishing for the first time. Our intuition is that the first published price cannot be affected by the current level of demand and is hence independent of it.

We prevent an omitted-variable bias by including the listings' location. Without controlling for location, the price could erroneously pick up geographical effects, hence also violating the second condition for a valid instrument.

Geographic location

Spatial location is a crucial determinant for success in the tourism industry, both in terms of proximity to the main attraction and in terms of economic clustering. Due to the importance of the 'sea and sun' segment for Balearic tourism, we computed the straight distance between each property and the closest beach, using GPS coordinates (using the Haversine formula, Robusto, 1957).

Secondly, we estimated geographical agglomeration using *spgwr*, a package of R-studio (Bivand *et al.*, 2007). For each property, we created a subsample with the closest properties[9] to identify a 'small' proximity area (what Assaf and Tsionas (2019) would define as a "tract") together with local performance metrics (i.e. local price and OR). The performance metrics of each cluster refer to June 2016 as they are expected to be a good predictor for the OR in the following months and also work as lagged variables for OR in July and August 2016. The geographically weighted statistics were computed using the tricube kernel function (Gollini *et al.*, 2013): the weight of each observation decreases with the increase in the distance from the listing under observation.

[9] Entire properties and rooms have been grouped separately.

Finally, local competition was proxied by the number of Airbnb listings active in July and August 2016 and within 1 km from each observation. The Haversine formula was also used to compute the straight distance between each property and all the local competitors.

The model specification

The generic model is specified as follows:

[Equation 4] $$y_1 = z\sigma + \alpha y_2 + u y_1$$

where y_1 is the OR, y_2 is the individual average price (EEV), and z is the 1xL vector of exogenous explanatory variables. We implemented a two-step procedure, first estimating EEV using a basic OLS and then estimating the demand model integrating the residuals from stage one[10].

To produce good approximations of confidence intervals (DiCiccio and Efron, 1996), we applied the bootstrap resampling method. The validity of the instruments was tested using a refutability test (Card, 1993), which is an alternative assessment of instrument excludability from the structural equation (French and Popovici, 2011). Results of the test proved the instruments 'suitability[11].

[10] Residuals were added in stage two as an additional variable.

[11] The characteristics that are required for an instrument to have statistical soundness were discussed in the Methodology section.

Results

Tables 3 and 4 present estimates of the ZOIB model on the entire sample (columns 1), on the geographical subsample of Mallorca (columns 2), Ibiza (columns 3) and on the listings published only by 'non-multiple hosts' (columns 4). This last subgroup includes the listings belonging to "*non-professional*" hosts (Li *et al.*, 2015) in a logic to exclude those listings managed by property management agencies. Table 3 shows the coefficients for the beta part while the zero-inflated part is presented in Table 4. In addition, column 5 of Table 3 presents estimates of the simple OLS model, without treating price endogeneity (as done in most of the previous literature).

[Insert Table 3 about here]

[Insert Table 4 about here]

The economic interpretation of the estimates is done through marginal analysis, and the following comments (which refer to the entire sample, columns 1 of Table 3 and 4, unless specified otherwise) should be read as the impact of a change in the specific regressor, other variables kept constant. We measure the marginal effects of the quantitative covariates at their median values and of the binary variables at their mean, preventing the description of unrealistic scenarios.

As expected from H1, price (AVERAGE_PRICE) has a negative effect on the OR (Table 3), but it is fundamental to highlight that we found no statistically significant effect of price on the probability of being unsold (Table 4). An increase of 10% in the price leads to a decrease of 0.0035 in the OR (coefficient -0.141, Table 3, column 1). Demand for Airbnb listings in the Balearic Islands proves to be strongly inelastic to price, much more than previously found in the literature (Li *et al*., 2015; Van der Borg *et al*., 2017; Gunter and Önder, 2018). An overestimation of the price effect on the OR would have occurred if the ZOIB model were run without treating for price endogeneity. The coefficient obtained (-0.401, associated with a marginal effect of -0.0084 when price changes by

10%) would be almost triple the one found in the basic model of column 1. These differences highlight the sensitivity of elasticity estimates to the econometric specification and clearly show that without properly treating endogeneity, elasticity might be strongly biased. The coefficient of AVERAGE_PRICE in the corresponding OLS model[12] (Table 3, column 5) implies an OR reduction of 2,9% when price increases by 10%. Even though this is not our preferred model, this coefficient suggests a more rigid demand for Airbnb listings in the Balearics compared to the case of Vienna (-0.518; Gunter and Önder, 2018) and Veneto Region (-0.373; Van der Borg *et al.*, 2017).

In line with H2, the price set by competitors influences the OR. If the price of the other listings is lower than the price set by the individual (DIFF_PRICE), the impact on OR is negative (-0.0151, Table 3, column 1). The coefficient was not statistically significant in the Mallorca sample (column 2), while DIFF_PRICE shows a significant negative effect in the Ibiza sample (-0.0209, column 3), hence confirming that competition must be understood at a very local level. This underlines the importance of correctly defining the market of relevance, as estimates are different, depending on the area considered in the analysis. Table 4 shows instead that setting a price higher than the average local price increases the probability of the listing being unsold (0.0783, Table 4, column 1), a result that is robust to the different subsamples.

In line with H3, the number of listings (N_LISTING) in the same area negatively affects individual performance (-0.0168, Table 3, column 1), a typical competition effect. There is no significant effect of N_LISTING on the probability of having OR=0 (Table 4) with an exception for Ibiza where the competition looks stronger (0.0285, Table 4, column 3).

Consistently with H4, we found a significant effect of eWOM. An increase in REVIEWS (in dozens) boosts the OR by 0.05 (marginal effect, computed on the coefficient estimated in Table 3, column 1)

[12] This model has a log-log specification (OR= ln (OR*100)) as was done in Gunter & Önder (2018). This specification makes easier to compare our results with the existing literature.

and decreases the probability of never being booked by 0.17 (marginal effect, computed on the coefficient estimated in Table 4, column 1). The volume of reviews also significantly affects the probability of being unsold (-0.833, Table 4, column 1). Listings with a visible score are associated with a higher OR. For listings with HIGH_RATE = 1, for example, the OR increases by 0.14 compared to listings NEVER_RATED (the base level, not included in the regression) and decreases the probability of being unsold by 0.25 (marginal effect). The impact of rating scores on OR and on the probability of being unsold is stronger for listings managed by 'non-multiple' hosts (Table 3 and Table 4, columns 4). These results are in line with existing literature (Viglia *et al*., 2016; Van der Borg *et al*., 2017; De Pelsmaker *et al*., 2018).

The econometric specification allows the evaluation of the impact of other characteristics linked to the listings' quality and to personal reputation. In line with H4, the number of pictures (PHOTOGRAPHS) generates a positive and significant effect on demand (0.022, Table 3, column 1), a likely consequence of the reduction in the information asymmetry between prospective guests and hosts. This result is coherent with Gunter and Önder (2018). The higher the number of pictures, the lower the probability of having OR=0 (-0.066, Table 4, column 1). As also found by Van der Borg *et al*., (2017), listings managed by a SUPER_HOST tend to have a higher OR (0.244, Table 3, column 1). The dummy 'SUPER_HOST' is not included in the zero-inflated part of the model since having a minimum number of bookings is mandatory for being a 'SUPER_HOST'.

Listings' LONGEVITY is (perhaps) surprisingly negatively associated with the OR. Any additional month of presence on Airbnb website produces a negative effect on the OR (-0.01, Table 3, column 1) and increases the probability of being unsold (0.018, Table 4, column 1).

The time lag between the booking request and the host's answer (RESPONSETIME) is negatively associated with OR (-0.0057, Table 3, column 1). This result is coherent with existing literature (Van der Borg *et al*., 2017; De Pelsmacker *et al.*, 2018). Variables proxying the dwelling size

(BEDROOM and MAXGUEST) are not robustly associated with OR. This result, which is partially in line with Gunter and Önder (2018), is not surprising, as price mostly captures the differences in size. Hosts setting INSTANT_BOOKING(s) are less likely to have OR=0 (-0.425, Table 4, column 1) while its effect on the OR is negative (-0.0940, Table 3, column 1). The only exception is for non-multiple hosts, where INSTANT_BOOKING has no statistically significant effect on OR (Table 3, column 4).

The property type does matter: ENTIRE_PROPERTY shows that apartments have on average higher ORs than shared and private rooms (0.262, Table 3, column 1), in line with previous research (Van der Borg, 2017). Higher values of MINIMUMSTAY (the minimum length of stay requested for booking) decrease the OR, while the request for a DEPOSIT boosts OR (0.0671, Table 3, column 1) perhaps because it is perceived as a signal of high quality.

In line with H5 and with existing works (Jeffrey *et al*., 2002; Van der Borg *et al*., 2017; Gunter and Önder, 2018), location is a critical driver of demand. First, compared to Mallorca (ISLAND_CODE=1), listings in Ibiza (ISLAND_CODE=2) tend to have higher ORs (0.149, Table 3, column 1)[13]. Being the Balearic Islands a destination mainly devoted to the sea and sun segment, and in line with H5, DISTANCE_BEACH is associated with a lower OR, although the effect is negligible: 1 km more of distance from the beach leads to a reduction of 0.002 in the OR (-0.00786, Table 3, column 1). The effect is statistically significant for Mallorca (-0.00524, Table 3, column 2) but not for Ibiza (Table 3, column 3). Data show that listings in Mallorca have an average distance to the beach of 4.67 km, while in Ibiza it is sensibly lower (1.80 km). This is tantamount to assuming that, in Ibiza, even farther properties are still in a limited distance from the main point of interest.

Finally, the local performance has a strong positive externality on individual performance, in line with H6 and with previous findings (Peiró-Signes *et al*., 2015). An increase (+0.1) in LOCAL_OR

[13] Menorca (ISLAND_CODE=3) has been excluded due to multicollinearity issues.

boosts the OR by +0.03 (0.0115, Table 3, column 1); being in a highly performing area guarantees a better OR (independent of the distance from the beach) and also decreases the probability of having OR=0 (-0.0108, Table 4, column 1).

Conclusions and Discussion

This paper studied the determinants of demand (measured through OR) for Airbnb listings in the Balearic Islands during the summer season of 2016. We built upon previous research on SE by introducing three important methodological and theoretical innovations in the analysis.

One, we applied the zero-inflated beta model to tackle the specificity of the OR distribution. The proposed model is an *ad hoc* tool for the study of proportions presenting a high share of zeros, as it is for OR in SE cases. A beta regression (Table 3) analysed the determinants of the positive part, when OR is in the range (0; 1), while a logit model studied the probability that the OR is equal to 0 (Table 4).

Two, in order to avoid biased estimates of price elasticity, we tackled endogeneity between OR and the price through a Control Function Approach and using an overidentified model. The inclusion of a proxy capturing 'location' effects prevented an omitted variable bias, which would lead to overestimating the negative relationship between price and OR and deliver biased estimates.

Three, we built the econometric specification in accordance with the theoretical framework of monopolistic competition, the relevant market regime for p2p accommodation.

The paper's results have a double reading. Empirically, they provide evidence for one of the most important tourism destinations worldwide, enabling the comparison with previous findings for other destinations. Methodologically, the paper shows how the adopted approach can tackle endogeneity and data issues by delivering quantitatively different and qualitatively superior estimates.

[Insert Table 5 about here]

Empirically, results are consistent with the model of monopolistic competition, the typical market structure for accommodation, and hence with hypotheses H1–H6 (Table 5). We found that the price negatively affects the OR of Airbnb listings in the Balearics (H1), although the estimated marginal effect is lower than values found in existing literature (Van der Borg *et al*. ,2017, Gunter and Önder, 2018). Moreover, the relative difference from the price of local competitors (H2) and the strength of market competition (H3), proxied by the number of Airbnb listings in the same area, significantly affect the OR.

In line with H4, online reputation is an important driver of demand: listings with bad and/or few reviews impact negatively on OR. The number of photographs, the host's responsiveness to clients' queries and the possibility of instant bookings are all important determinants of OR linked to the host's reputational capital. In a market characterised by asymmetric information, hosts need to put a considerable effort into self-marketing and develop specific personal branding strategies. Moreover, location confirmed its crucial role for the success of a listing, something that was overlooked by previous empirical analyses. *Ceteris paribus*, being in highly-performance area (H6) and/or close to tourist attractions (H5) is associated with a higher OR.

Methodologically, our contribution highlights the importance of choosing the appropriate model, as coefficients are highly sensitive to misspecifications. The coefficient of price triples when running the ZOIB model without controlling for price endogeneity. Similarly, running a simple OLS model, (Table 3, column 5) replicating Van Der Borg *et al*. (2017) and Gunter and Önder (2018), leads to a much higher sensitivity of OR to price. On the contrary, when controlling for price endogeneity and for the high share of zeros in the dependent variable (Table 3, column 1), demand appears more inelastic. The price, although still a statistically significant determinant of the OR, has only a negligible marginal effect: OR decreases only of -0.0035 if price is increased by 10%.

Moreover, this methodology allows disentangling the pure effect of price on OR from more general considerations about why some properties are not rented at all. As reported in Table 4 most of the explanatory variables show coefficients in line with the continuous part of the model (Table 3) but, surprisingly, we found no statistically significant relationship between price and the probability of being unrented. Apparently, never rented properties have some peculiar characteristics that make them undesirable for tourists, regardless of price.

Our general conclusion is that the quantitative impact of explanatory variables on OR is therefore dependent on the estimation techniques. The correct methodology adopted in this study shows that the role of price as a determinant of sales is strongly reduced. It's not all about price, indeed.

Recommendations for managers

Wrong models might lead to wrong results. Practitioners should be aware that relying on studies adopting an incorrect methodology could engender ineffective pricing strategies and negatively impact on business performance. Prospective hosts should be aware that OR could be less sensitive to price variations than expected. Pricing strategies determined by the standard model of the literature (Table 3, Column 5) would imply that a 10% price increase would lead to a 2.9 % decrease in the OR, at the median level for the covariates. Once endogeneity is correctly accounted for (marginal effects computed on Table 3, column 1), the same price change would decrease OR by only 0.6%. In this regard price has mainly a strategic role, while its tactical use is barely effective (differently from Abrate and Viglia, 2016): hosts should mainly check that prices are in line with those of local competitors and with the listing's characteristics. More important that price is the adoption of suitable booking management tools and, above all, investing in personal branding and reputation (consistent with Mauri *et al*., 2018).

Location is another key driver of demand. The geographical position of a dwelling cannot be adjusted, particularly if the property/room is the one where the host lives. Refurbishments undertaken to stimulate demand should reflect the profitability of such investment, since projects in 'unattractive' areas could be at loss. Table 6 shows a naïve simulation (built on strong simplifications and assumptions) with the sole aim of emphasizing how location (column 1) is predominant in determining the performance. By considering the selling price per square metre[14] (column 2), the price for an 80-square-metre apartment (column 3), the Airbnb average seasonal revenue (column 4) and the yearly payment of a mortgage (column 5)[15], it arises the heterogeneity of the investment returns (column 6). Moreover, location also impacts on the probability of being unrented: an apartment in the centre of Palma de Mallorca has an estimated probability of being unrented in the whole season of approximately 4%, while an apartment in Campos (located on the south side of Mallorca), *ceteris paribus*, has an estimated probability of 15%[16].

[Insert Table 6 about here]

Limitations and further research

The present study is not free of limitations. One, albeit an important tourism destination, data cover a relatively short period and are restricted to the case of Balearic Islands. The elasticities and the relationships direction might change for different destinations or types of tourism, something that should be investigated in a comparative framework.

Two, in the current study we did not account for seasonality since the analysis is only performed during the high season 2016. Tourism demand tends to be less elastic in peak season and it could be interesting to compare price sensitivity across seasons.

Three, we neglected the role of special events as determinants of occupancy rates, mainly because daily data were aggregated over the period under consideration, transforming the panel into a cross-section. Future extensions should include special events as determinant factors of OR in a panel framework.

[14] IBESTAT 'precios y viviendas', available at: https://ibestat.caib.es/ibestat/estadistiques/economia/construccio-habitatge/preu-habitatge

[15] We assume a fixed interest of 2% and a mortgage time of 20 years, with monthly payments.

[16] Marginal analysis was computed using the value of the regressors at their median value.

Four, although the OR is a key indicator, it is likely that hosts aim at maximizing REVPAR more than occupancy and, in this sense, research should also consider the use of dynamic pricing and its impact on overall performance.

Five, hosts can take advantage of multiple platforms and rent via different distribution channels. This represents a crucial veil of ignorance for research: in fact, it cannot be determined whether a listing ‘blocked’ on Airbnb was instead booked through a different channel or was used instead by the owners. The understanding of why a listing is unavailable on Airbnb requires further in-depth analysis.

References

Abrate, G. and Viglia, G. (2016), “Strategic and tactical price decisions in hotel revenue management”, *Tourism Management*, Vol. 55, pp. 123-32.

Assaf, A. G. and Tsionas, M. G. (2019), “Forecasting occupancy rate with Bayesian compression methods”, *Annals of Tourism Research*, Volume 75, pp. 439-49.

Bivand, R., Yu, D., Bivand, M. R., Brunsdon, M. C. and Fortheringham, S. (2007), “The spgwr Package”, *R Foundation for Statistical Computing*.

Buis, M. (2012). ZOIB: Stata module to fit a zero-one inflated beta distribution by maximum likelihood.

Card, D. (1993). *Using geographic variation in college proximity to estimate the return to schooling* (No. w4483). National Bureau of Economic Research.

De Pelsmacker, P., Van Tilburg, S. and Holthof, C. (2018), “Digital marketing strategies, online reviews and hotel performance” *International Journal of Hospitality Management*, Vol. 72, pp. 47-55.

DiCiccio, T. J. and Efron, B. (1996), “Bootstrap confidence intervals”, *Statistical Science,* Vol. 11 No. 3, pp. 189-212.

Ferrari, S. and Cribari-Neto, F. (2004), “Beta regression for modelling rates and proportions”, *Journal of Applied Statistics*, Vol. 31 No. 7, pp. 799-815.

French, M. T. and Popovici, I. (2011), “That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research”, *Health Economics*, Vol. 20 No. 2, pp. 127-46.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C. and Harris, P. (2013), “GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models”, *arXiv*:1306.0413.

Gunter, U. and Önder, I. (2018), “Determinants of Airbnb demand in Vienna and their implications for the traditional accommodation industry”, *Tourism Economics*, Vol. 24 No. 3, pp. 270-93.

Hu, N., Zhang, J. and Pavlou, P. A. (2009), “Overcoming the J-shaped distribution of product reviews”, *Communications of the ACM*, Vol. 52 No. 10, pp. 144-47.

Jeffrey, D., Barden, R. R., Buckley, P. J. and Hubbard, N. J. (2002), “What makes for a successful hotel? Insights on hotel management following 15 years of hotel occupancy analysis in England”, *Service Industries Journal*, Vol. 22 No. 2, pp. 73-88.

Kim, Y. and Uysal, M. (1997), “The endogenous nature of price variables in tourism demand studies”, *Tourism Analysis*, Vol. 2 No. 1, pp. 9-16.

Lei, W. S. C. and Lam, C. C. C. (2015), “Determinants of hotel occupancy rate in a Chinese gaming destination”, *Journal of Hospitality and Tourism Management*, Vol. 22, pp. 1-9.

Li, J., Moreno, A., & Zhang, D. J. (2015). Agent behavior in the sharing economy: Evidence from Airbnb. *Ross School of Business Working Paper Series*, *1298*, 2015.

Liang, S., Schuckert, M., Law, R. and Chen, C. C. (2017), “Be a “Superhost”: the importance of badge systems for peer-to-peer rental accommodations”, *Tourism Management*, Vol. 60, pp. 454-65.

Liang, L. J., Choi, H. C. and Joppe, M. (2018). “Understanding repurchase intention of Airbnb consumers: perceived authenticity, electronic word-of-mouth, and price sensitivity”, *Journal of Travel & Tourism Marketing*, Vol. 35 No. 1, pp. 73-89.

Lockyer, T. (2005), “The perceived importance of price as one hotel selection dimension”, *Tourism Management*, Vol. 26 No. 4, pp. 529-37.

Lurkin, V., Garrow, L. A., Higgins, M. J., Newman, J. P. and Schyns, M. (2017), “Accounting for price endogeneity in airline itinerary choice models: an application to continental US markets”, *Transportation Research Part A: Policy and Practice*, Vol. 100, pp. 228-46.

Mauri, A. G. and Minazzi, R. (2013), “Web reviews influence on expectations and purchasing intentions of hotel potential customers”, *International Journal of Hospitality Management*, Vol. 34, pp. 99-107.

Mauri, A. G., Minazzi, R., Nieto-García, M. and Viglia, G. (2018), “Humanize your business. the role of personal reputation in the sharing economy”, *International Journal of Hospitality Management*, Vol. 73, pp. 36-43.

Mumbower, S., Garrow, L. A. and Higgins, M. J. (2014), “Estimating flight-level price elasticities using online airline data: a first step toward integrating pricing, demand, and revenue optimization”, *Transportation Research Part A: Policy and Practice*, Vol. 66, pp. 196-212.

Öğüt, H. and Onur Taş, B. K. (2012), “The influence of internet customer reviews on the online sales and prices in hotel industry”, *The Service Industries Journal*, Vol. 32 No. 2, pp. 197-214.

Oh, H. (2003), “Price fairness and its asymmetric effects on overall price, quality, and value judgments: the case of an upscale hotel”, *Tourism Management*, Vol. 24 No. 4, pp. 387-99.

Ospina, R. and Ferrari, S. L. (2012), “A general class of zero-or-one inflated beta regression models”, *Computational Statistics & Data Analysis*, Vol. 56 No. 6, pp. 1609-23.

Peiró-Signes, A., Segarra-Oña, M. D. V., Miret-Pastor, L. and Verma, R. (2015), “The effect of tourism clusters on US hotel performance”, *Cornell Hospitality Quarterly*, Vol. 56 No. 2, pp. 155-67.

Pekgün, P., Griffin, P. M., & Keskinocak, P. (2013). *An empirical study for estimating price elasticities in the travel industry*. Working Paper at University of South Carolina.< http://www2. ie. psu. edu/griffin/price. pdf>(accessed 08.02. 14).

Robusto, C. C. (1957). The cosine-haversine formula. *The American Mathematical Monthly*, *64*(1), 38-40.

Porter, M. E. (1998). *Clusters and the new economics of competition* (Vol. 76, No. 6, pp. 77-90). Boston: Harvard Business Review.

Salop, S. C. (1979), “Monopolistic competition with outside goods”, *Bell Journal of Economics*, Vol. 10 No. 1, pp. 141-56.

Teubner, T., Saade, N., Hawlitschek, F. and Weinhardt, C. (2016), “Its only pixels, badges, and stars: on the economic value of reputation on Airbnb” in *Australian Conference on Information Systems.*

Tussyadiah, I. P. (2016), “Factors of satisfaction and intention to use peer-to-peer accommodation”, *International Journal of Hospitality Management*, Vol. 55, pp. 70-80.

Tussyadiah, I. P. and Pesonen, J. (2016), “Impacts of peer-to-peer accommodation use on travel patterns”, *Journal of Travel Research*, Vol. 55 No. 8, pp. 1022-40.

Van der Borg, J., Camatti, N., Bertocchi, D., & Albarea, A. (2017). The Rise of the Sharing Economy in Tourism: Exploring Airbnb Attributes for the Veneto Region. *University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No, 5*.

Viglia, G., Minazzi, R., Buhalis, D., (2016). “The influence of e-word-of-mouth on hotel occupancy rate”, *International Journal of Contemporary Hospitality Management*, Vol. 28 No. 9, pp. 2035-51.

Viglia, G., Mauri, A. and Carricano, M. (2016), “The exploration of hotel reference prices under dynamic pricing scenarios and different forms of competition”, *International Journal of Hospitality Management*, Vol. 52, pp. 46-55.

Wooldridge, J. M. (2011), “Fractional response models with endogeneous explanatory variables and heterogeneity”, *CHI11 Stata Conference* (No. 12). Stata Users Group.

Wooldridge, J. M. (2015), “Control function methods in applied econometrics”, *Journal of Human Resources*, Vol. 50 No. 2, pp. 420-45.

Yang, Y., Mao, Z. and Tang, J. (2017), “Understanding guest satisfaction with urban hotel location”, *Journal of Travel Research*, Vol. 57 No. 2, pp. 243-59.

Ye, Q., Law, R. and Gu, B. (2009), “The impact of online user reviews on hotel room sales”, *International Journal of Hospitality Management*, Vol. 28 No. 1, pp. 180-82.

Zekan, B., Önder, I., & Gunter, U. (2019). Benchmarking of Airbnb listings: How competitive is the sharing economy sector of European cities?. *Tourism Economics*, *25*(7), 1029-1046.

Table 1- A summary of the literature

Reference	Subject	Methodology	Case study	Results
Assaf, A. G., & Tsionas, M. G. (2019)	Hotel	Nonlinear compressed Vector Autoregression	USA	(1) The hotel's OR is strongly affected by the other hotels' prices within the same area ("tract"); (2) Compressed VAR outperforms existing models.
De Pelsmacker, P., Van Tilburg, S., & Holthof, C. (2018)	Hotel	Mediation Model	Flanders (Belgium)	(1) Review volume positively affects OR; (2) Digital market strategies positively affect OR; (3) The effect of reviews and market strategies is stronger for chain hotels; (4) The effect of reviews and market strategies increases with the number of stars.
Gunter, U., & Önder, I. (2018)	Airbnb	Cluster-robust OLS	Vienna (Austria)	(1) Price negatively affects OR, although demand is price inelastic; (2) Quicker responses to booking requests lead to higher OR; (3) Listing size positively affects OR; (4) The number of photos positively affects OR; (5) The distance from the city centre negatively affects OR.
Jeffrey, D., Barden, R. R., Buckley, P. J., & Hubbard, N. J. (2002)	Hotel	Time Series and OLS	England (UK)	(1) Location affects OR; (2) A romantic setting increases OR; (3) Conference facilities increase OR; (4) Management practices strongly influence OR.
Lei, W. S. C., & Lam, C. C. C. (2015)	Hotel	OLS	Macao (China)	(1) Price negatively affects OR, although demand is price inelastic; (2) Casino facilities negatively affect OR; (3) Hotel star rating negatively affects OR; (4) The number of tourism arrivals positively affects OR; (5) The total number of available rooms positively affects OR.
Li, J., Moreno, A., & Zhang, D. J. (2015)	Airbnb	OLS	Chicago (USA)	(1) Professional hosts (hosts managing more than one listing) achieve higher OR; (2) Price inefficiency leads to lower OR.
Van der Borg, J., Camatti, N., Bertocchi, D., & Albarea, A. (2017).	Airbnb	OLS	Veneto region (Italy)	(1) Price negatively affects OR, although demand is price inelastic; (2) Entire properties have higher OR than private or shared rooms; (3) Location significantly affects OR; (4) A lower response time to the booking request leads to a higher OR; (5) Super hosts achieve higher OR; (6) The overall rating positively affects OR.
Viglia, G., Minazzi, R., & Buhalis, D. (2016).	Hotel	OLS	Rome (Italy)	(1) An increase in the review score positively affects OR; (2) The review volume has a positive effect on OR but with diminishing returns.
Our study	Airbnb	Zero-Inflated beta model	Balearic Islands (Spain)	(1) Price negatively affects OR, although demand is highly inelastic; (2) Location and reputation strongly affect OR; (3) Being in a high-performing area positively affects OR; (4) Price does not explain why some listings are never booked; (5) Overlooking price endogeneity leads to biased estimates.

Table 2- Descriptive statistics

Variable	Description	Mean/Proportion	Std. Dev.	Min	Max
average_price	natural logarithm of the average price (regardless of the booking status), at time *t*	5.308	0.887	2.398	9.294
bedrooms	number of bedrooms	2.329	1.484	0	10
deposit	deposit required (Dummy variable)	0.566	0.496	0	1
diff_price	difference between individual and local average price, at time t_{-1}	9.164	456.85	-1635.920	9966.979
distance_beach	distance to the closest beach (in km)	3.397	4.485	0.007	23.300
entire_property	entire home/apartment (Dummy variable)	0.839	0.368	0	1
high_rate	score rating (4.5; 5] (Dummy variable)	0.469	0.499	0	1
instant_booking	offer instant booking (Dummy variable)	0.228	0.419	0	1
iscode	Island code (1=Mallorca; 2=Ibiza; 3=Menorca; 4=Formentera)	1.505	0.709	1	4
local_OR	local average occupancy rate at time t_{-1}	0.264	0.096	0	0.827
longevity	number of months since listing creation	17.084	13.584	0	85.833
low_rate	score rating [0; 4] (Dummy variable)	0.047	0.211	0	1
maxguest	max number of people who can be accommodated	5.476	2.987	1	16
medium_rate	score rating (4; 4.5] (Dummy variable)	0.145	0.352	0	1
minimumstay	minimum number of nights required per booking	3.810	2.052	1	21
n_listing	number of listings in the same area	324.170	449.950	0	1530
never_rated	the listing has never been rated (Dummy variable)	0.190	0.393	0	1
OR	average occupancy rate at time *t*	0.488	0.346	0	1
photographs	number of photographs	23.656	15.208	1	469
published_daily	natural logarithm of the price set per night (by default)	5.151	0.865	2.398	9.258
published_monthly	natural logarithm of the price set per month (by default)	8.265	0.894	4.942	12.590
published_weekly	natural logarithm of the price set per week (by default)	6.925	0.903	4.304	12.287
responsetime	time elapsed between guest request and host answer (in hours)	4.462	6.304	0	24
reviews	number of reviews (in dozens)	12.853	18.413	0	224
super_host	Host is a "super host" (Dummy variable)	0.073	0.261	0	1
y2res	residual from step 1 of CFA	0.000	0.302	-3.222	1.602

time t= bimester July-August 2016   time t_{-1}= June 2016

Table 3. Determinants of OR when 0 < OR ≤ 1, ZOIB model, beta part

	(1) OR whole sample	(2) OR Mallorca island	(3) OR Ibiza island	(4) OR Non-multiple hosts	(5) OR OLS estimator, whole sample
proportion					
average_price	-0.141***	-0.323***	-0.222***	-0.376***	-0.291***
	[-3.91]	[-8.64]	[-3.99]	[-6.22]	[-17.79]
y2res	-0.391***	-0.258***	-0.307***	0.066	
	[-9.25]	[-5.62]	[-4.58]	[0.75]	
diff_price	-0.015***	0.007	-0.021***	-0.009	-0.006*
	[-2.94]	[0.77]	[-3.26]	[-0.91]	[-1.77]
n_listing	-0.017***	-0.025***	-0.011***	-0.003	-0.013***
	[-6.83]	[-7.41]	[-2.72]	[-0.63]	[-7.11]
reviews	0.204***	0.162***	0.339***	0.152***	0.080***
	[23.41]	[17.16]	[17.04]	[10.23]	[15.52]
high_rate	0.591***	0.543***	0.604***	0.615***	0.438***
	[25.00]	[20.17]	[12.40]	[11.42]	[24.93]
medium_rate	0.514***	0.460***	0.505***	0.605***	0.366***
	[17.49]	[13.50]	[8.89]	[9.19]	[16.72]
low_rate	0.379***	0.357***	0.370***	0.287***	0.271***
	[8.85]	[6.54]	[5.18]	[3.05]	[8.35]
photographs	0.022***	0.029***	0.018	0.052***	0.023***
	[3.57]	[3.92]	[1.63]	[4.34]	[5.25]
super_host	0.244***	0.264***	0.197***	0.166***	0.169***
	[7.31]	[6.87]	[2.88]	[3.07]	[6.96]
longevity	-0.010***	-0.010***	-0.013***	-0.009***	-0.006***
	[-14.57]	[-11.72]	[-9.40]	[-6.81]	[-11.75]
responsetime	-0.006***	-0.006***	-0.005*	-0.008***	-0.001
	[-3.81]	[-3.37]	[-1.81]	[-2.74]	[-1.01]
bedrooms	-0.006	-0.014	0.079***	0.019	0.033***
	[-0.42]	[-0.78]	[2.97]	[0.68]	[3.06]
maxguests	-0.028***	-0.014	-0.033***	-0.011	-0.012**
	[-3.71]	[-1.38]	[-2.70]	[-0.76]	[-2.12]
instantbooking	-0.094***	-0.070***	-0.155***	0.031	-0.082***
	[-4.69]	[-3.06]	[-3.46]	[0.83]	[-5.31]
entireproperty	0.262***	0.462***	0.213***	0.499***	0.301***
	[6.91]	[10.13]	[3.43]	[7.54]	[12.05]
minimumstay	-0.041***	-0.048***	-0.010	-0.013	-0.029***
	[-8.79]	[-8.53]	[-1.10]	[-1.56]	[-8.90]
deposit	0.067***	0.041**	0.105***	0.056*	0.082***
	[3.74]	[1.99]	[2.70]	[1.67]	[5.96]
Ibiza	0.149***			0.227***	0.228***
	[4.62]			[4.02]	[11.39]
Formentera	0.010			0.208*	0.141***
	[0.15]			[1.68]	[2.93]
distancebeach	-0.008***	-0.005***	-0.014	-0.006*	-0.007***
	[-4.15]	[-2.64]	[-1.35]	[-1.75]	[-5.09]
local_OR	0.012***	0.014***	0.010***	0.005**	0.008***
	[9.48]	[9.05]	[3.96]	[2.30]	[8.60]
_cons	0.387**	1.096***	0.750***	1.421***	0.052
	[2.56]	[6.54]	[2.81]	[5.50]	[0.69]
Oneinflate					
_cons	-3.377***	-3.520***	-3.166***	-2.968***	
	[-71.91]	[-59.18]	[-38.36]	[-39.27]	
N	16392	11243	4802	4103	14201
Wald chi2	5238.64	4052.55	1233.43	1049.50	
Prob<chi2	0	0	0	0	
Adj.R^2					0.297

Table 4. Determinants of OR when OR = 0, ZOIB model, zero-inflated part

	(1) OR whole sample	(2) OR Mallorca island	(3) OR Ibiza island	(4) OR Non-multiple hosts
Zeroinflated				
average_price	-0.055	-0.150	-0.164	0.036
	[-0.55]	[-1.33]	[-1.26]	[0.16]
y2res	0.080	-0.047	0.673***	0.393
	[0.65]	[-0.32]	[3.94]	[1.18]
diff_price	0.078***	0.093***	0.079***	0.085***
	[8.09]	[4.95]	[6.72]	[3.53]
n_listing	0.011	0.017	0.029**	-0.006
	[1.31]	[1.25]	[2.50]	[-0.27]
reviews	-0.883***	-0.653***	-1.442***	-0.539***
	[-12.83]	[-8.62]	[-10.04]	[-5.12]
high_rate	-1.045***	-0.961***	-1.246***	-1.429***
	[-14.55]	[-10.62]	[-10.16]	[-8.57]
medium_rate	-1.094***	-0.839***	-1.387***	-1.567***
	[-11.13]	[-6.87]	[-8.36]	[-6.07]
low_rate	-0.845***	-0.591***	-0.977***	-1.393***
	[-6.44]	[-3.25]	[-5.00]	[-3.47]
photographs	-0.066***	-0.062**	-0.077**	-0.012
	[-3.12]	[-2.09]	[-2.40]	[-0.26]
longevity	0.018***	0.016***	0.03***	0.031***
	[8.40]	[5.76]	[7.71]	[6.44]
responsetime	0.017***	0.023***	0.012**	0.020**
	[4.54]	[4.57]	[2.08]	[2.45]
bedrooms	0.011	0.128**	-0.113*	-0.028
	[0.26]	[2.13]	[-1.85]	[-0.28]
maxguests	-0.014	-0.072**	0.066**	-0.028
	[-0.67]	[-2.21]	[2.17]	[-0.55]
instantbooking	-0.425***	-0.490***	-0.159	-0.318*
	[-5.70]	[-5.35]	[-1.14]	[-1.66]
entireproperty	-0.292***	-0.626***	0.142	-0.463*
	[-2.62]	[-4.37]	[0.84]	[-1.77]
minimumstay	-0.004	-0.041**	0.069***	0.077***
	[-0.35]	[-2.39]	[3.51]	[2.76]
deposit	-0.009	0.124*	-0.287***	-0.074
	[-0.17]	[1.77]	[-2.95]	[-0.54]
Ibiza	0.936***			0.454**
	[10.37]			[2.17]
Formentera	0.783***			0.663*
	[4.12]			[1.65]
distancebeach	-0.002	-0.001	0.039	0.021
	[-0.28]	[-0.08]	[1.62]	[1.47]
local_OR	-0.011***	-0.015***	-0.028***	-0.009
	[-2.77]	[-2.94]	[-4.06]	[-0.91]
_cons	-0.538	0.369	0.686	-1.342
	[-1.27]	[0.72]	[1.06]	[-1.40]
ln_phi				
_cons	1.172***	1.170***	1.224***	1.432***
	[106.57]	[90.08]	[56.80]	[64.47]
N	16392	11243	4802	4103
Wald chi2	5238.64	4052.55	1233.43	1049.50
Prob<chi2	0	0	0	0

Table 5. Hypotheses and summary of empirical evidence

Hp.	Factor	Expected Sign	Empirical support	Discussion
H1	Price	Negative	Yes	Compared to Gunter & Önder (2018), Lei & Lam (2015), Van der Borg *et al*. (2015) we found that demand is much more rigid, although price negatively affects demand.
H2*	(*p*-*P*)	Negative	Yes	We support Assaf & Tsionas (2019) in finding that price of competitors is important, and we highlight the local dimension of competition.
H3	Number of Airbnb listings	Negative	Yes	According to Salop model (1979), individual demand is negatively related to the number of competitors.
H4	Reviews	Positive	Yes	Similar to De Pelsmacker *et al*. (2018), Van der Borg *et al*. (2015), Viglia *et al*. (2016) we found a positive effect of online reputation on demand.
H5	Distance to the point of interest	Negative	Yes	Similar to Gunter & Önder (2018), Jeffrey *et al*. (2002), Van der Borg *et al*. (2015) we found that location is a significant determinant of individual demand. Again, we highlight local idiosincracies.
H6	Local performance	Positive	Yes	We support Peiró-Signed *et al.* (2015) in finding that being in a highly performing (geographical) cluster has a positive impact on demand.

* p = individual price; P = price set by competitors

Table 6. Expected profitability of short-term rental via Airbnb

Municipality [1]	Sell Price €/m2 [2]	Apartment price [3]	Airbnb summer revenue* [4]	Yearly payment amount [5]	Expected economic result
Calvià	2.506,70 €	200.536,00 €	11.794,98 €	12.173,74 €	-378,76 €
Ciutadella	1.516,10 €	121.288,00 €	7.932,86 €	7.362,91 €	569,95 €
Eivissa	2.772,40 €	221.792,00 €	25.559,66 €	13.464,00 €	12.095,66 €
Inca	1.105,90 €	88.472,00 €	6.401,54 €	5.370,78 €	1.030,76 €
Maó	1.261,60 €	100.928,00 €	8.871,70 €	6.126,00 €	2.745,70 €
Marratxí	1.652,90 €	132.232,00 €	8.069,23 €	8.027,28 €	41,95 €
Palma	1.612,50 €	129.000,00 €	12.348,82 €	7.831,07 €	4.517,75 €

* This scenario relies on strong simplifications: (i) Airbnb prices and ORs are stable over the years; (ii) listings are rented only during the period between May and September

Figure 1. The occupancy rate of Airbnb listings in the Balearic Islands