Quality & Quantity (2020) 54:1633–1653
https://doi.org/10.1007/s11135-019-00925-2

Identifying factors affecting the status of superhost: evidence from Sardinia and Sicily

Giulia Contu[1] · **Claudio Conversano**[1] · **Luca Frigau**[1] · **Francesco Mola**[1]

Published online: 8 August 2019

Abstract

Airbnb is currently one of the most developed new forms of hospitality. It is essentially an online platform that connects the owners of apartments or rooms with potential guests. The core of the Airbnb model is the host, who offers a support to guests and favor a link between the guests and the tourism destination. Airbnb awards the best hosts with the badge of "Superhost", which is attributed according to four elements: the occupancy rate, the number of reservations, the response rate and the cancellation policy. This paper focuses on hosts and their activities. Specifically, the main goal is understanding if the four aforementioned elements actually influence the attribution of the "Superhost" badge to hosts operating in two of the main Italian touristic destinations: Sardinia and Sicily. Furthermore, the link between the four basic elements and other "managerial factors" is analyzed. Logistic and Probit models are used for these purposes and the main findings are derived from the computation of marginal effects. The results show a direct impact of the four Airbnb variables and of other "managerial" variables, as for instance the presence of extra fees for cleaning or similar services, on the probability to be a superhost.

Keywords Airbnb · Superhost · Logistic regression · Probit regression · Marginal effects

1 Introduction

Airbnb is a company operating as an online marketplace for peer-to-peer accommodation rental services. It is one of the most famous online platforms where it is possible to book single rooms or whole apartments for one or more days. Founded in 2008 in San Francisco, in the last few years it has grown significantly, becoming an important tool for tourists who look for a different, often less expensive kind of accommodation (Guttentag 2015; Aznar et al. 2017). Actually, it operates in more than 65,000 cities and 191 countries and it sells millions of room nights for tourists around the globe. In 2016, Airbnb has declared

✉ Claudio Conversano
conversa@unica.it
http://people.unica.it/claudioconversano

[1] Department of Business and Economics, University of Cagliari, Viale Frà Ignazio 17, 09045 Cagliari, Italy

160 million guests. In the same year, more than three millions of announcements were published.

The number of hosts is 2.9 millions and an average of 800,000 overnight stays are spent each day in rooms or apartments booked on Airbnb.[1] The number of people that subscribe as hosts increases at a rate of around 14,000 new hosts per month.[2] To become a host is only necessary to decide to share a proper space, a room or an entire apartment, and to enroll on the Airbnb platform. Moreover, before to appear in the platform, it is necessary to decide, and to declare, specific aspects related to the host activity: a minimum and maximum number of nights, how far in the future guests can book, cleaning fees, weekly discounts or eventually special prices and so on.[3]

The role of hosts is at the core of the Airbnb activity. This role is so relevant that Airbnb awards the best hosts with the status of "superhost". In fact, each 3 months, Airbnb attributes this badge to new hosts and confirms or deletes the status of previous superhosts. The status is defined taking into account four different criteria. Airbnb declares that to be a superhost it is necessary, in the previous 12 months:

1. to host at least 10 stays a year;
2. to honor every reservation unless there is an extenuating circumstance;
3. to respond to customer requests within 24 h at least 90% of the time;
4. to achieve a 4.8+ overall rating.

The last criterion has been recently modified. Until June 2018, it was necessary to have a five-star rating from at least 80% of accommodated guests. In other words, hosts must be active, reliable, responsive and highly rated.[4] However, it is not clear if these criteria have a direct influence on the status of superhost, and if all the variables are equally important, or if one variable is more important than the others.

Recently, a study carried out on Airbnb hosts in San Francisco (US) has demonstrated that the four Airbnb criteria influence the attribution of the superhost badge and that it is possible to identify a rank among these criteria (Gunter 2018). In particular, this study revealed that the most important criterion is rating, followed by a reliable cancellation behavior, responsiveness and a sufficient demand. The author of Gunter (2018) has recognized three different limitations of his study. The first is a geographical limitation, as only the San Francisco area is analyzed. The second is a time limitation: data are focused only on 2016. The third limitation is generated by data themselves, as it used scraped data and not official data.

The research findings presented in Gunter (2018) are the starting point of the analysis presented in this paper, whose main aim is to overcome two of the above-mentioned limitations: the geographical one and that deriving from the use of non official data. We focus on the two most important and attractive Italian islands, Sardinia and Sicily, and analyze a dataset obtained from the company that officially manages the Airbnb's data. Thus, it does not make use of scraped data. Specifically, the main goal of this study is twofold. First, it is designed to comprehend if the Airbnb criteria can really influence the status of superhost and if the rank defined by Gunter in Gunter (2018) can be successfully applied in different

1 https://www.airbnb.com/host/homes?from_footer=1&locale=en, accessed on 15/02/2019.
2 https://www.airbnb.com/host/homes?from_footer=1&locale=en, accessed on 15/02/2019.
3 https://www.airbnb.com/b/setup, accessed on 15/02/2019.
4 https://www.airbnb.com/Superhost?locale=en, accessed on 30/11/2018.

geographic areas. Knowing which variables are more relevant can actually support hosts improving their chances to become a superhost. Second, we want to assess if the status of superhost not only depends from several factors considered individually, as it has been investigated in previous studies, but if these factors interact one with another to improve the chance for a host to become a superhost. To accomplish for this second goal, we aim to comprehend if other variables are involved in the attribution of the superhost badge. Specifically, the influence of variables related to the management decision, i.e. those strictly related to the management of the Airbnb activity, has been considered. To our knowledge, this is the first study that investigates the crucial elements that influence the status of superhost taking into account not only the main variables used to attribute the badge, but also the variables related to the management of the host activity.

To sum up, the effort of this study is to answer the following questions:

Q_1 = Are the variables that Airbnb declares to use in the identification of superhost really related to the status of superhost for hosts operating in Sardinia or Sicily? Is it possible to identify a rank among these variables?

Q_2 = Can other variables, and the interactions among them, influence the status of superhost for hosts operating in Sardinia or Sicily?

We apply both the Logistic and the Probit model to answer the two above-mentioned research questions. The results underline that the statements made in Gunter (2018) can be extended straightforwardly to Sardinia and Sicily. Moreover, it is discovered that other variables, and their joint effect, can increase the probability of being a superhost.

Five sections compose this study. The first Section is focused on the review of the literature, the second on the presentation of an overview of Logistic and Probit regression and marginal effects as well as of both data and descriptive statistics. The results of the application of the two models on Sardinia and Sicily are shown in the fourth Section. Finally, the last Section is focused on the managerial implications and on some concluding remarks.

2 Literature review

In the last years, different researchers have studied Airbnb, focusing on different aspects.

Some researchers have studied Airbnb taking into account the *development of the sharing economy*, as Airbnb is based on a platform that connects individuals with resources that they want to share to individuals that need those resources (Li et al. 2015; McNamara 2015). This exchange between hosts and guests has determined the acknowledgment of Airbnb as one of the best-known sharing economy companies (Zervas et al. 2015; Gutt and Herrmann 2015; Gutierrez et al. 2016; Quattrone et al. 2016; Lutz and Newlands 2018; Gibbs et al. 2018; Roelofsen and Minca 2018; Liang et al. 2017; Dudás et al. 2017; Ert et al. 2016; Gunter 2018). For instance, Zervas et al. (2015, p. 2) have defined Airbnb as a center piece of the so-called sharing economy; Dudás et al. (2017) have defined it as one of the peer-to-peer based online platforms that promote user-generated content, sharing, and collaboration. However, other researchers, as for instance Oskam and Boswijk (2016), have criticized the vision of Airbnb as a part of sharing economy because of the monetary nature of Airbnb. They have stated that the sharing economy is based on a simple interaction between two persons and not on a monetary exchange.

Other scholars have analyzed Airbnb w.r.t. the *trustworthiness and perceived trust*. For instance, Zhang et al. (2018) have focused their study on the perceived trust, establishing that the trust is based on different aspects as rating scores, textual reviews, photos, the badge of superhost, accuracy and completeness of the information provided. Moreover, Zervas et al. (2015) have analyzed the importance of ratings and reviews. The same results have been highlighted in Guttentag and Smith (2017), where it is reported that Airbnb's trust is based on different aspects as, for instance the reviews, the price, and the photos. Other researchers have recognized the importance of using the photos to promote the accommodation and to produce an impact on the guests' decisions (Ert et al. 2016). These studies are relevant because they allow us to identify the fundamental elements for the definition of the Airbnb service and for the creation of trust in guests.

Other studies have focused on *consumer segmentation*. Lutz and Newlands (2018) have analyzed this aspect in the effort to comprehend whether Airbnb users prefer to book a shared room or an entire home. The results suggest that those who travel alone or with friends prefer to stay in shared rooms, whereas guests that travel with parents or children prefer to chose entire apartments. Moreover, the choice is influenced also by socio-economic status: guests with a higher socio-economic status are more likely to book an entire home.

Another important topic analyzed has been the choice of the *price for the Airbnb accommodation*. Some researchers, as for instance Wang and Nicolau (2017) and Gibbs et al. (2018), have attempted to identify which variables can have a direct impact on the price of Airbnb accommodations.

Additionally, Choi et al. (2015), Zervas et al. (2017), Guttentag and Smith (2017), Martin-Fuentes et al. (2018) and Blal et al. (2018) have analyzed the *impact of Airbnb on the hotels' activity*. Not all studies show the same results. For instance, Guttentag and Smith have stated that *many Airbnb guests use the service in place of a hotel, and especially mid-range hotels* (Guttentag and Smith 2017, p. 9). This statement seems to contradict the above results of Choi et al. that have suggested that Airbnb has no effect on hotel revenue in a specific area as Korea (Choi et al. 2015).

Another considered aspect is the *localization of Airbnb accommodation* in the territories and the impact of Airbnb on cities (Gurran and Phibbs 2017). Dudás et al. (2017) have defined a method to map the spatial distribution of Airbnb accommodations, whereas Gutiérrez et al. (2017) have analyzed their spatial distribution. Quattrone et al. (2016) have mapped the Airbnb accommodations in London to comprehend the socio-economic characteristics of areas, and to evaluate the existence of differences between Airbnb listings of rooms and those of entire houses.

Finally, some researchers have focused their attention on *the badge of superhost*. Liang et al. (2017) have analyzed how the status of superhost can influence published reviews and the behavior of ratings. Specifically, they have discovered that the accommodations managed by a superhost are more likely to receive reviews than those managed by a normal host, that the price of the accommodation is negatively associated with the review volume, and that guests prefer to spend more money for accommodations with the superhost badge. In the same perspective, Roelofsen and Minca have analyzed the phenomenon of Airbnb studying the importance and the pervasiveness of the biopolitical spatialities in the sharing economies and particularly in the Airbnb platform. They have analyzed the capacity of Airbnb to create communities and interactions between guests and hosts. Particular attention is posed on the study of superhosts, defined as *the champions of the Airbnb world of hospitality* and as *stellar human beings who excel in performing hospitality at their home* (Roelofsen and Minca 2018, p. 177). These scholars have also underlined the importance of becoming superhosts in terms of a bigger visibility on the Airbnb platform. Wang and Nicolau (2017) have also discovered a positive relationship between the superhost status and the price, evidencing an increase in the income for the hosts that obtain the badge.

Additionally, Ma et al. (2017) have analyzed how the hosts describe themselves on their Airbnb profile pages and the perceived trustworthiness of their profiles. They have discovered that superhosts have significantly more extensive host profile. Within this framework, as evidenced before, an interesting analysis has been realized by Gunter (2018) in order to comprehend if the variables used by Airbnb to define the status of superhost can really impact on the probability to be a superhost. He has applied the Logit and Probit model with the aim to evaluate which variable can influence this probability. Moreover, the use of marginal effects has allowed identifying the impact on this probability.

As previously stated, the present paper builds on the study presented in Gunter (2018) and tries to define a suitable statistical model that might be operationalized to understand the basic factor determining the shift from the status of host to the superhost one. The importance of studying the superhost status comes from the fact that the badge allows hosts to obtain several benefits. Airbnb summarizes the benefits in four main areas.[5] First of all, a bigger visibility: superhosts are placed better in the online researches, so they can be found more easily. Moreover, a superhost filter is located in the Airbnb website improving, another time, the possibility to be found more easily and quickly. Finally, a badge is added to Superhosts' profiles to certify their high standards, a guarantee of high quality services for Airbnb guests. The second advantage is related to the exclusive *peaks*, as called by Airbnb. These peaks are defined through a recognition of an extra referral bonus, the possibility to choose between a 100$ travel coupon or one free professional photo session in available markets and the application of discounts on next products and on Airbnb experience. The third advantage is related to the insider access. Airbnb provides a priority support when a superhost has a problem in his activity. Additionally, Airbnb recognizes an early access to new features offering the possibility to be the first to develop new programs and to test features before they are launched to the broader host community. Finally, superhosts can participate in exclusive events and product launches organized by Airbnb. The fourth advantage is related to the tools of the host business. Airbnb offers a tax service that allows obtaining a special discount and resources from the Airbnb tax partners. Moreover, Airbnb states in its website[6] that superhosts earn 22% more than other hosts. Improved income can be considered as a fifth, certainly not negligible advantage.

3 Research methodology

3.1 Basics of Logistic and Probit models

In order to answer to the two research questions, we estimate the probability of being a superhost through both the *Logistic model* and the *Probit model*. These models have been chosen since they allow to reach the main goal of inferencing the effects of the covariates over the binary response variable *Superhost*. Furthermore, they evaluate the marginal effects, providing more interpretable results especially for the coefficients of discrete covariates (Powers and Xie 2008).

Let us consider the matrix $\mathbf{X}' = \left(X_0, X_1, X_2, \dots, X_p\right)$, where $X_0 = (1, \dots, 1)'$ is a constant variable of length n and the other elements are the covariates expressed by a collection of p

[5] https://www.airbnb.com/superhost.

[6] https://www.airbnb.com/Superhost?locale=en, accessed on 30/11/2018.

independent variables of length n. The response variable $Y = (y_1, \dots, y_n)'$ is dichotomous: it is equal to 1 when the host is a *Superhost* and to 0 if the host is not a *Superhost*. The conditional probability for a superhost is

$$\Pr(Y = 1|\mathbf{X}, \boldsymbol{\beta}) = g(\mathbf{X}'\boldsymbol{\beta}) = \pi(\mathbf{X}) \tag{1}$$

The vector $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)$ is an unknown vector of $p + 1$ regression coefficients. Several functions $g(\cdot)$ can be considered to estimate the relationship between Y and $\mathbf{X}$. In the generalized linear modeling framework, the function $g(\cdot)$ is known as *link function* (see, for example (Fahrmeir and Tutz 1994)). Hereafter, we focus on logit and probit functions.

The logit-link function used in the logistic regression model is specified as

$$g(\mathbf{X}'\boldsymbol{\beta}) = \frac{e^{\sum_{i=0}^{p} \beta_i X_i}}{1 + e^{\sum_{i=0}^{p} \beta_i X_i}} \tag{2}$$

The probability $\pi(\mathbf{X})$ as defined in Eq. (1) is related to $\boldsymbol{\beta}$ through $g^{-1}(\mathbf{X}'\boldsymbol{\beta})$, consequently

$$\log\left(\frac{\pi(\mathbf{X})}{1 - \pi(\mathbf{X})}\right) = \sum_{i=0}^{p} \beta_i X_i \tag{3}$$

As for the probit model, instead, the probit-link function is specified as

$$g(\mathbf{X}'\boldsymbol{\beta}) = \boldsymbol{\Phi}(\mathbf{X}'\boldsymbol{\beta}) \int_{-\infty}^{\mathbf{X}'\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du, \tag{4}$$

and $\boldsymbol{\Phi}(\cdot)$ is the cumulative distribution function of a standard normal distribution. The probit transformation, given by the inverse of the cumulative distribution function of a standard normal distribution, i.e. $g^{-1}(\mathbf{X}'\boldsymbol{\beta})$, allows us to relate $\pi(\mathbf{X})$ to $\boldsymbol{\beta}$. Consequently:

$$\boldsymbol{\Phi}^{-1}(\pi(\mathbf{X})) = \sum_{i=0}^{p} \beta_i X_i \tag{5}$$

Both the logistic and probit model parameters are estimated through maximum likelihood. The likelihood function is

$$\mathcal{L}(\boldsymbol{\beta}|Y, \mathbf{X}) = \prod_{i=1}^{n} \left(g(\mathbf{X}'\boldsymbol{\beta})^{y_i}[1 - g(\mathbf{X}'\boldsymbol{\beta})]^{(1-y_i)}\right) \tag{6}$$

Maximum likelihood estimation consists in finding the coefficient vector $\boldsymbol{\beta}$ maximizing $\log[\mathcal{L}(\boldsymbol{\beta}|Y, \mathbf{X})]$.

As shown above, the main difference between logistic and probit is the link function, although both models generally give similar results. From a mathematical point of view, the logistic distribution has slightly fatter tails: the variance of a logistically distributed random variable is about $\pi^2/3$, whereas that of a (standard) normally distributed variable is one. That is to say, the conditional probability $\Pr(Y = 1|\mathbf{X}, \boldsymbol{\beta})$ approaches 0 or 1 at a slower rate in logit than in probit. But in practice there is no compelling reason to choose one over the other. For these reasons, we consider both approaches in the estimation of the probability of being a superhost.

Table 1 Airbnb data

Variables	Means
Superhost	1 = superhost and 0 = otherwise
Cancellation policy	1 = cancellation policy and 0 = no cancellation policy
Cleaning fee	1 = cleaning fee and 0 = no cleaning fee
Number of bookings	Number of unique reservations in the last 12 months
Response rate	The percentage of time a host responds to potential guests within 24 h
Overall rating	Overall customer rating from on a scale of 1–5
Max guests	The maximum number of guests the listing can accommodate
Extra people fee	Charge $N/A = no$ additional fee
Instant book enabled	Yes = the property can be booked without any host/guest communication
Occupancy rate	The percentage of : (days with a reservation) / (total number of days with a reservation and available). Calculation excludes blocked days
Latitude	
Longitude	

For both models (logistic and probit), a marginal effect is defined as *the rate at which y changes at a given point in covariate space, with respect to one covariate dimension and holding all covariate values constant* (Leeper 2017, p. 7). A marginal effect is computed through partial derivatives. For the kth the independent variable, in the case of continuous variables a unit change of x_{ik} determines a β_k change in y_i. Consequently, β_k can be considered the marginal effect of x_{ik}

$$\frac{\partial y_i}{\partial x_{ik}} = \beta_k \tag{7}$$

For a discrete independent variable the marginal effect is given by

$$E(y_i|x_{ik} = 1) - E(y_i|x_{ik} = 0) = \beta_k \tag{8}$$

Furthermore, a marginal effect can also be computed taking into account the conditional probability of $y_i = 1$. Specifically,

$$\frac{\partial \Pr(y_i = 1|\mathbf{x}_i)}{\partial x_{ik}} = \frac{\partial g(\mathbf{x}_i'\boldsymbol{\beta})}{\partial x_{ik}} = f(\mathbf{x}_i'\boldsymbol{\beta})\beta_k \tag{9}$$

where $f(\cdot)$ denotes the density function. This quantity is the rate of change in the success probability in the neighborhood of a particular value of x.

3.2 Data and descriptive statistics

The study has been realized using a dataset provided by *Airdna*®, a company that deals with Airbnb data. The Airdna dataset includes information concerning different aspects as for instance the hosts and their activities, the revenues, the number of reservation days, the characteristics of Airbnb accommodations, the judgments on the apartments and the coordinates of the accommodations.

The variables we are going to consider in our analysis concern the information related to the host and to aspects involving his management activity, as shown in Table 1.

The variable Superhost is chosen as a response variable. It is a binary variable that assumes value equal to one when the host is a superhost, and zero otherwise.

We consider four Airbnb variables as covariates in a first step of the analysis, that are used to specify a model to answer the research question Q_1. They are:

- the *Cancellation policy*, which has been transformed into a binary variable. It assumes value zero if the host choses a flexible cancellation policy and one if he opts for a strict or moderate policy;
- the *Cleaning fee*, it assumes value one if the host requires a cleaning fee and zero otherwise;
- the number of reservations (*Number of Bookings*);
- the rate of a host responding to his guests (*Response Rate*);
- the value that the guests express to judge the whole Airbnb experience (*Overall Rating*).

In a second step, other variables are included in the model. These variables can be classified as managerial variables, as they are related to the managerial decisions of the host. Specifically, the subset is composed by the following variables:

- *Max Guests*, the number of maximum guests that can be hosted;
- *Extra People fee*, the price to pay to add one or more persons to the reservation;
- *Instantbook Enabled*, a service that allows guests to book an accommodation without an explicit host approval, facilitating the reservation process.
- *Occupancy rate*, the proportion of days in a year when the accommodation is booked. This rate is related with the number of bookings, which is one of the variables that Airbnb uses to assign the status of superhost. This variable cannot be controlled or determined directly by the host.

These variables are taken into account because, as evidenced before, they are important elements for a host in order to develop trust and to improve his communication ability. They are considered as fundamental aspects to become a superhost and are thus used to specify the model to answer the research question Q_2.

The empirical analysis is made on data concerning reservations made during 2016 in Sardinia and Sicily. In the original dataset, 24,651 Airbnb hosts were located in Sardinia and 42,436 in Sicily. The original dataset has been cleaned removing observations presenting missing values at least in one of the variables *Superhost*, *Cancellation policy*, *Number of Bookings*, *Response Rate* and *Overall Rating*. These variables are considered in the literature (e.g. Guttentag and Smith 2017; Gunter 2018) as the most important ones in determining the status of superhost. At the same time, we considered as outliers the observations that lie beyond the 95% whiskers of the boxplot of the variables *Number of Bookings*, *Response Rate* and *Overall Rating*, on both sides. These observations were removed from the original dataset. Following this data cleaning process, the final dataset is composed of 19,756 hosts, of which 11,086 are located in Sicily and 8670 in Sardinia.

The accommodations of the two islands considered in the final dataset are represented in Figs. 1 and 2. It is worth noting that they are mostly located near the coast. This aspect confirms what stated in Quattrone et al. (2016) and Boros et al. (2018): Airbnb listings tend to be concentrated near popular touristic destinations and in touristic areas.

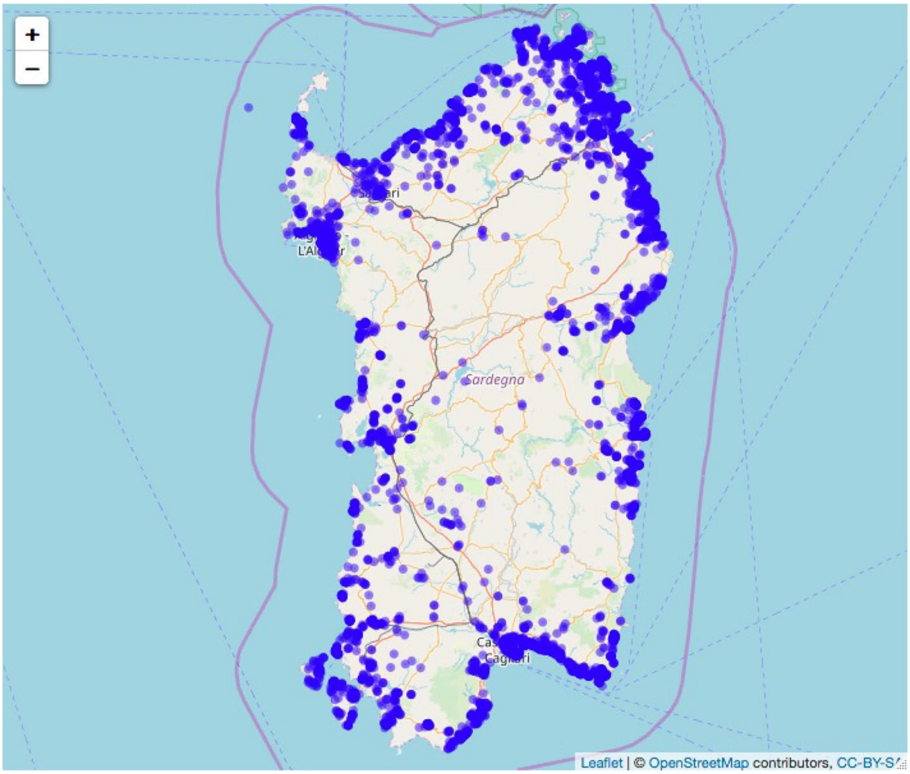

Fig. 1 Airbnb in Sardinia

Tables 2 and 3 report the descriptive statistics for the observed variables for the two groups of Airbnb hosts.

From Table 2 it is possible to notice that the two regions have the same average levels of occupancy and response rates, maximum number of guests and overall rating. One important difference emerges w.r.t. the number of bookings: in Sicily, the mean number of bookings in a year equals 9.71 (median = 7.00). This average level appears as relevantly higher than that observed in Sardinia, where the mean (median) of the number of bookings is 6.71 (5.00). For all the observed numerical variables, there is more variability observed for Sicilian hosts: the coefficient of variation (last column of Table 2) observed for Sicily is always higher than that observed for Sardinia.

Table 3 reports the observed proportions of occurrences for the binary variables, including the variable Superhost that is used as response in the regression models. It is possible to notice that the proportion of superhosts is very high and rather similar in the two regions (about 0.90 for Sicily and about 0.93 for Sardinia). Important differences emerge, instead, for other variables: Sicilian hosts tend more to apply a moderate/strict cancellation policy as well as to require a cleaning fee to their clients. Sardinian hosts tend more to require additional fees for people who is added next to the reservation and are more orientated towards instant bookings.

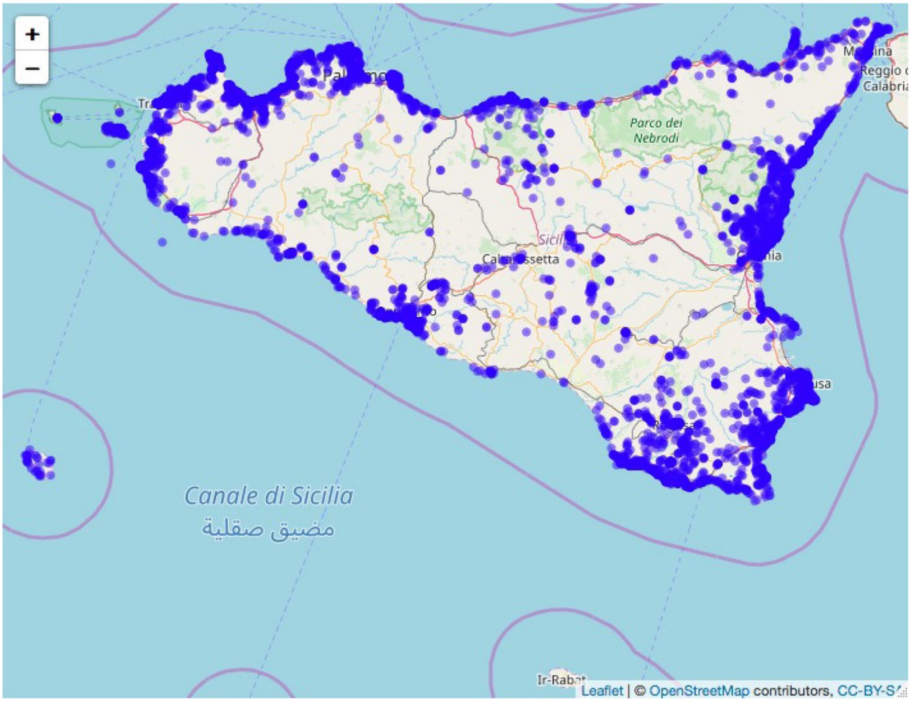

Fig. 2 Airbnb in Sicily

Table 2 Descriptive statistics of the observed numerical variables: minimum (min), maximum (max), median, mean, standard deviation (SD) and coefficient of variation (coef.var)

Variable	Min	Max	Median	Mean	SD	Coef.var
Occupancy rate	0.00	0.66	0.17	0.21	0.16	0.76
Max guests	1.00	16.00	4.00	4.53	2.43	0.54
Number of bookings	0.00	43.00	7.00	9.71	9.06	0.93
Response rate	76.00	100.00	100.00	98.86	3.83	0.04
Overall rating	3.60	5.00	4.70	4.62	0.36	0.08
Occupancy rate	0.00	0.63	0.19	0.22	0.15	0.66
Max guests	1.00	16.00	4.00	4.45	2.08	0.47
Number of bookings	1.00	25.00	5.00	6.71	5.17	0.77
Response rate	71.00	100.00	100.00	98.20	5.11	0.05
Overall rating	3.50	5.00	4.70	4.58	0.39	0.09

In the next Sections, we will see that all the differences observed in Tables 2 and 3 influence the outcomes of the regression models.

Table 3 Proportion of occurrences for the observed binary variables (p_1 is the observed proportion for Airbnb hosts operating in Sicily whilst p_2 is observed proportions for Airbnb hosts operating in Sardinia)

p	Variable				
	Superhost	Cancellation policy	Extra people fee	Cleaning fee	Instant-book enabled
p_1	0.899	0.329	0.363	0.424	0.650
p_2	0.932	0.262	0.558	0.341	0.755

4 Results and implications

The Logistic and Probit models have been applied on the Sardinia and Sicily data in order to understand if the four major Airbnb variables identified in Gunter (2018) have a direct impact on the probability to obtain the superhost badge (first research question, Q_1) as well as to investigate about the role of other important variables as well as on their interaction effects (second research question, Q_2).

4.1 The importance of the four major variables and their possible ranking

The four major variables used by Airbnb, specifically overall rating, cancellation policy, number of bookings and response rate, have a direct influence on the probability to be a superhost, as evidenced in Table 4. Results obtained from the two types of models provide the same information and are consistent one with another. Furthermore, the results obtained for the two islands also underline the existence of a rank among the variables: a rank that is confirmed by the results of the two models and by the computation of the marginal effects (bottom panel of Table 4). These results are in line with the statements made in Gunter (2018). The most important variable is overall rating, followed by cancellation policy, number of bookings and response rate.

It must be underlined that the variable "overall rating" is also able to capture the general judgement on the experience and the satisfaction level. It is therefore the synthesis of different aspects, as for instance the neatness of the accommodation, its location, etc. Additionally, this variable is really important because it also gives a measure of the quality of the service. This element is essential for a host that wants to become a superhost and wants to maintain this position. Moreover, as evidenced in Gutt and Herrmann (2015), this positive judgement can also impact on the rent price and improve the earnings rates of the host.

The second variable in the ranking is *Cancellation Policy*. A strict cancellation policy is considered as a warranty for the guest (Gunter 2018). It is supposed that the accuracy required from the host is also required by the host. Specifically, it is supposed that the choice of a strict cancellation policy is a signal of accuracy and honesty that the host requests to his guests. The same accuracy and honesty, on the other side, will be requested to the host, that is supposed not to cancel the reservation. When the guest books an accommodation, he certainly hopes to spend his holiday pleasantly and without problems. The cancellation of a room or an apartment is therefore a relevant problem, so in order to become a superhost it is very important not to delete reservations.

Table 4 Logistic and Probit model: estimated coefficients and marginal effects

	Estimated model							
	Logistic regression				Probit regression			
Variable	$\hat{\beta}$	$\hat{\sigma}_{\beta}$	z	Pr(> |z|)	$\hat{\beta}$	$\hat{\sigma}_{\beta}$	z	Pr(> |z|)
	Sicily							
(Intercept)	− 21.335	1.361	− 15.67	0.00	− 10.683	0.661	− 16.15	0.00
N of bookings	0.059	0.003	18.73	0.00	0.031	0.002	17.78	0.00
Response rate	0.034	0.012	2.98	0.00	0.017	0.006	3.02	0.00
Overall rating	3.128	0.153	20.51	0.00	1.531	0.072	21.15	0.00
Cancellation policy	0.328	0.074	4.42	0.00	0.178	0.039	4.60	0.00
	Sardinia							
(Intercept)	− 20.628	1.514	− 13.63	0.00	− 9.664	0.691	− 13.98	0.00
N of bookings	0.122	0.007	17.11	0.00	0.061	0.004	16.04	0.00
Response rate	0.035	0.012	2.88	0.00	0.016	0.006	2.92	0.00
Overall rating	2.829	0.192	14.75	0.00	1.271	0.086	14.87	0.00
Cancellation policy	0.292	0.108	2.71	0.01	0.139	0.053	2.62	0.01
	Marginal effects							
	Logistic regression				Probit regression			
Variable	*AME*	$\hat{\sigma}_{AME}$	z	Pr(> |z|)	*AME*	$\hat{\sigma}_{AME}$	z	Pr(> |z|)
	Sicily							
N of bookings	0.0048	0.00	18.93	0.00	0.0048	0.00	17.94	0.00
Response rate	0.0028	0.00	2.98	0.00	0.0027	0.00	3.02	0.00
Overall rating	0.2576	0.01	20.02	0.00	0.2379	0.01	20.96	0.00
Cancellation policy	0.0259	0.01	4.61	0.00	0.0267	0.01	4.78	0.00
	Sardinia							
N of bookings	0.0071	0.00	16.46	0.00	0.0070	0.00	15.62	0.00
Response rate	0.0020	0.00	2.87	0.00	0.0019	0.00	2.91	0.00
Overall rating	0.1636	0.01	13.94	0.00	0.1467	0.01	14.31	0.00
Cancellation policy	0.0160	0.01	2.86	0.00	0.0154	0.01	2.75	0.01

Number of booking and *Response Rate* are ranked in the third and fourth position, respectively. These two variables show roughly the same values. In fact, the value of marginal effects related to the two variables are very similar both in Siciliy and in Sardinia.

To be a superhost, it is necessary to host at least 10 stays during the previous year. A high number of reservations is a signal that the host is probably offering a service highly appreciated by guests. In Liang et al. (2017) it is also showed that the superhost badge attracts more bookings.

To sum up, in order to become a superhost it is necessary to have a high number of reservations. Since we are observing destinations (Sardinia and Sicily) characterized by a seasonal nature of tourism where tourists use to stay often for a couple of weeks in the period from May to September, being a superhost in those regions implies that it is very likely that the accommodation is often booked out of season also.

At the same time, as underlined before, a high number of reservations guarantees the conservation of the superhost status. It is like a virtuous circle, where the number of reservations is the key element that allows both obtaining the superhost status and to maintain it through time. Moreover, an elevate number of bookings can be also considered as a distinctive element of a host that works "professionally". The label *professional* has been used by several researchers to describe Airbnb hosts. The first to use of this term occurred in Li et al. (2015) that classified the Airbnb hosts as *professional* or *unprofessional*. Specifically, the label *professional* was used for the hosts that rent more than one property. On the contrary, the label *unprofessional* was used for the hosts that offer only one property on the Airbnb platform. The same label is applied in this study. However, in this case the label *professional* is used to identify those hosts that have the status of Superhost. On the contrary, the label *unprofessional* is used for those hosts that do not have the badge of Superhost. The separation of professional (super)hosts from unprofessional (non-super)hosts ones operated in this study is motivated by the relevant effort realized by the Superhost to reach the four above-mentioned criteria required by Airbnb. Only an accurate management of the Airbnb activity and a constant effort to improve the quality of the service allow to obtain the badge and to maintain it. This effort is a clear signal that the aim of a host is to market a profitable product, and therefore we use it as an indication that the superhost is actually a professional host.

The variable Response Rate has the same relevance. This variable is directly calculated by the platform and it measures how fast the host responds to guest's online enquiry (Roelofsen and Minca 2018). This aspect is relevant for three reasons. Firstly, the Airbnb activity is based on the creation of a relationship between hosts and guests. Secondly, the choice of an Airbnb accommodation is based (or at least should be based) on the desire to meet new people and to live the holiday differently. Finally, a host that answers quickly changes completely the travel experience, as it supports and accompaign guests along their entire vacations.

The four variables have generally the same impact and assume roughly the same values on the two different Italian islands. Thus, it is possible to answer Q_1 stating that the variables that Airbnb declares to use in the identification of superhost are really related to the status of superhost for hosts operating in Sardinia or Sicily and a rank among these variables exists.

Interesting differences are found on the value of the variables *Number of Bookings*, *Overall rating* and *Cancellation Policy*. The marginal effect of *Cancellation Policy* and *Overall Rating* is higher in Sicily whilst that of *Number of Bookings* is higher in Sardinia. These differences are observed for both the logistic and the probit models. The values obtained for the two islands are similar to the results obtained by Gunter (2018) for San Francisco. This means that the same model is therefore applicable for other geographic areas.

4.2 The role of "managerial variables" and the interaction effects

In a second step of the analysis, after assessing the direct influence that the four Airbnb variables have on the status of superhost, we attempted to comprehend if other variables can influence directly the probability to be a Superhost. It is clear that only the four variables are used by Airbnb to attribute this badge. However, it is important to establish if other aspects can influence this probability in order to support hosts in their effort to improve the quality of the service they offer and their position in the Airbnb platform. At the same time, it is interesting to investigate whether these additional variables have the same effect on the probability of being a superhost in two different areas.

It has been shown that aspects such as services, rental rules and customer reviews can influence the possibility to obtain the badge of Superhost (Wang and Nicolau 2017). We focus on the impact that the variables related to the management of an Airbnb accommodation have on the probability to become a superhost. The Airbnb platform, in fact, gives hosts the possibility to make different choices concerning the reservation process and the accommodation rules, that can change completely not only the booking experience but also the entire stay. For instance, they can ask guests to pay a security deposit, fix a minimum or maximum number of days, and choose what services they want to offer.

Specifically, in our analysis we focus on the following additional variables that we define "managerial variables": *Occupancy Rate*; *Max Guests*; *Extra People Fee*; *Cleaning Fee* and *Instant book Enabled*. The first two variables are quantitative, whereas the remaining three are qualitative.

The variable *Occupancy rate* is intended as a proxy of the quality of the service. It is not directly controllable by the host but strictly related to the quality of the service offered by the host since it depends on the proportion of positive reviews made by guests. Consequently, it can be considered as a variable controlled by the host. Moreover, it has been demonstrated that guests generally write positive rather than negative reviews, and that they normally share only the good experiences they had during their stays, minimizing the bad aspects (Bridges and Vásquez 2018).

Therefore, if a host gets many reviews, they will be positive in most cases, and a high number of positive reviews has a positive impact also on the probability to become a superhost. As a corollary, a high number of positive reviews also generates trust in potential guests, and this affects even more positively the occupancy rate and thus the possibility to obtain the superhost badge.

The Logistic and Probit models have been applied in order to comprehend the direct impact of the managerial variables on the probability to be a superhost. Beside the main effects part of the model, we searched for significant cross-product interaction terms of both the second and third order to understand if the basic variables and the managerial ones interact one with another and in different ways to determine the superhost status. Results of this second step of the analysis are summarized in Table 5.

We considered the main effects model as our starting model and added all the possible interaction terms of both the second and third order. At the end, to define the final model we removed the interaction terms that were not significant. Of course, the final model has been defined taking into account of the well known hierarchy principle: a third-order interaction effect is kept in the model only if all the interactions of the second order involving the same covariates are included as well. Interestingly and rather surprisingly, the final model obtained for the two regions include the same main effects and the same interaction terms. Nevertheless, important differences arise w.r.t. the estimated coefficients and the marginal effects.

From Table 5, it is possible to notice that the estimated coefficients for the four variables used in the first step of the analysis (whose results are reported in Table 4), but not for *Number of Bookings*, have all the same positive effect on the probability of being a superhost. This result characterizes models estimated for the two islands. The number of bookings, if considered alone, has a negative effect on the estimated probability, particularly for hosts operating in Sicily. The other managerial variables included in the main effects part of the model lead to different findings for the two areas: *Extra People Fee* is the

Table 5 Logistic and Probit model: estimated coefficients of the main effects model and the interaction terms

	Estimated model							
	Logistic regression				Probit regression			
Variable	$\hat{\beta}$	$\hat{\sigma}_\beta$	z	Pr(> |z|)	$\hat{\beta}$	$\hat{\sigma}_\beta$	z	Pr(> |z|)
	Sicily							
(Intercept)	− 20.635	4.25	− 4.85	0.00	− 9.21	1.88	− 4.89	0.00
N of bookings	− 0.626	0.12	− 5.31	0.00	− 0.38	0.06	− 6.40	0.00
Response rate	0.110	0.04	2.79	0.01	0.05	0.02	2.77	0.01
Overall rating	1.405	0.33	4.21	0.00	0.58	0.15	3.84	0.00
Cancellation policy	0.251	0.08	3.24	0.00	0.13	0.04	3.25	0.00
Occupancy rate	18.261	10.65	1.71	0.09	7.10	5.34	1.33	0.18
Max guests	− 0.102	0.02	− 5.94	0.00	− 0.05	0.01	− 6.08	0.00
Extra people fee	0.412	0.08	5.38	0.00	0.21	0.04	5.36	0.00
Cleaning fee	16.250	4.90	3.32	0.00	7.11	2.25	3.16	0.00
Instantbook enabled	− 0.080	0.07	− 1.11	0.26	− 0.04	0.04	− 1.00	0.32
N of bookings × overall rating	0.137	0.02	5.59	0.00	0.08	0.01	6.70	0.00
N of bookings × cleaning fee	− 0.34	0.17	− 2.02	0.04	− 0.18	0.09	− 2.13	0.03
Overall rating × cleaning fee	− 0.159	0.43	− 0.37	0.71	− 0.04	0.20	− 0.21	0.83
Occupancy rate × response rate	− 0.162	0.11	− 1.51	0.13	− 0.06	0.05	− 1.13	0.26
Occupancy rate × cleaning fee	− 43.634	14.48	− 3.01	0.00	− 19.95	7.38	− 2.70	0.01
Response rate × cleaning fee	− 0.154	0.04	− 3.46	0.00	− 0.07	0.02	− 3.37	0.00
N of bookings × overall rating × cleaning Fee	0.072	0.03	2.05	0.04	0.04	0.02	2.16	0.03
Occupancy rate × response rate × cleaning Fee	0.437	0.15	3.00	0.00	0.20	0.07	2.69	0.01
	Sardinia							
(Intercept)	− 10.356	4.26	− 2.43	0.01	− 4.35	1.87	− 2.33	0.02
N of bookings	− 0.685	0.23	− 3.01	0.00	− 0.41	0.11	− 3.76	0.00
Response rate	0.037	0.38	0.71	0.05	0.00	0.02	0.21	0.83
Overall rating	1.228	0.47	2.59	0.01	0.48	0.21	2.28	0.02

Table 5 (continued)

Variable	Estimated model							
	Logistic regression				Probit regression			
	$\hat{\beta}$	$\hat{\sigma}_{\beta}$	z	Pr(> |z|)	$\hat{\beta}$	$\hat{\sigma}_{\beta}$	z	Pr(> |z|)
Cancellation policy	0.231	0.11	2.05	0.04	0.12	0.06	2.13	0.03
Occupancy rate	− 13.746	14.71	− 0.93	0.35	− 8.02	6.90	− 1.16	0.24
Max guests	− 0.121	0.03	− 4.52	0.00	− 0.06	0.01	− 4.45	0.00
Extra people fee	0.424	0.09	4.50	0.00	0.20	0.05	4.21	0.00
Cleaning fee	− 0.674	5.58	− 0.12	0.90	− 0.31	2.46	− 0.13	0.90
Instantbook enabled	− 0.211	0.11	− 1.94	0.05	− 0.10	0.05	− 1.89	0.06
N of bookings × overall rating	0.156	0.05	3.29	0.00	0.09	0.02	4.07	0.00
N of bookings × cleaning fee	− 0.968	0.34	− 2.85	0.00	− 0.49	0.16	− 3.01	0.00
Overall rating × cleaning fee	− 0.832	0.61	− 1.36	0.17	− 0.43	0.25	− 1.61	0.11
Occupancy rate × response rate	0.156	0.15	1.05	0.29	0.09	0.07	1.28	0.20
Occupancy rate × cleaning fee	28.691	17.29	1.66	0.10	14.79	8.22	1.80	0.07
Response rate × cleaning fee	0.044	0.04	0.91	0.36	0.02	0.02	1.04	0.30
N of bookings × overall rating × cleaning fee	0.215	0.07	3.04	0.00	0.11	0.03	3.20	0.00
Occupancy rate × response rate × cleaning fee	− 0.299	0.17	− 1.72	0.09	− 0.15	0.08	− 1.86	0.06

unique variable that present a positive sign of the estimated coefficient for both islands and this coefficient is highly significant. Conversely, *Occupancy Rate* and *Cleaning Fee* have a positive effect on the superhost status in Sicily only. These two variables have no important effect for Sardinian hosts if considered alone. For these hosts, it seems that the presence of the instant-booking option is an important element that influences the estimated probability. The analysis of the main effects part of the model also shows that the maximum number of guest has a positive effect on the superhost status for hosts located in Sicily and a negative effect for those operating in Sardinia.

Moving to the second-order interactions, the output reported in Table 5 shows that number of bookings and overall rating have together a positive effect on the probability of being a superhost in both regions, whilst the opposite result is obtained for the joint effect of the number of bookings and the presence of a cleaning fee. Those results seem to suggest that the turnover of guests and the rating are important elements to improve the quality of the service, but in accommodations with many reservations where additional fees are requested, as for example for cleaning, are not straightforwardly linked to the superhost stauts, as the required extra-payment is not always perceived by guests as an improvement of the service.

It could be guessed that this effect is particularly important for Sicily if we consider that the joint effects of the response rate and the cleaning fee, and of the occupancy rate and the cleaning fee, since both cause the probability of being a superhost to decrease.

The differences observed for the second-order interaction effects are reflected in the two interactions of the third order that characterize the models estimated for Sicily and Sardinia. They show that increasing at the same time the number of bookings and the overall rate causes the probability of being a superhost to increase at most for hosts who require a cleaning fee: this effect is very important in Sardinia as the value for the estimated coefficient in both models are higher than those obtained for Sicilian hosts. In Sicily, the probability of being a superhost is highly influenced by the joint effect of occupancy rate and response rate: this means that hosts replying quickly at clients' requests are very likely to increase their occupancy rates and this effect is more pronounced in the case a cleaning fee is also required. Conversely, there is a negative joint effect of these three variables on the probability of being a superhost for Sardinian hosts: in these cases, it seems that these additional services are not particularly appreciated.

For the models estimated in this second stage of the analysis, marginal effects are not reported to save space. Those related to the main effects part of the model and, in particular, to the added "managerial variables" are consistent with the values of estimated coefficients. Those concerning the interaction effects cannot be computed. Nevertheless, the latter can be represented graphically as they help in understanding how the joint effect of two or more covariates influence the probability of the occurrence of a certain event.

In this respect, we focus on third-order interactions and represent in Fig. 3 the two terms obtained from our second model separately for Sicily and Sardinia. The two scatterplots in the first row of Fig. 3 represent the marginal effect of overall rating and the number of bookings on the probability of being a superhost for a host operating in Sicily (left panel) or in Sardinia (right panel). The interaction effect involves the possibility to ask for a cleaning fee also, as we represent with two separate lines hosts requiring this additional fee (dotted line) and those who do not require it (solid line). It is possible to notice that the number of bookings and the overall rating jointly have a direct influence on the superhost status in both regions, but this influence is more marked for hosts requiring the cleaning fee. In Sardinia, if the number of bookings is below 7.5 and/or the overall rate is below 0.20 this effect is not so evident, to be exact.

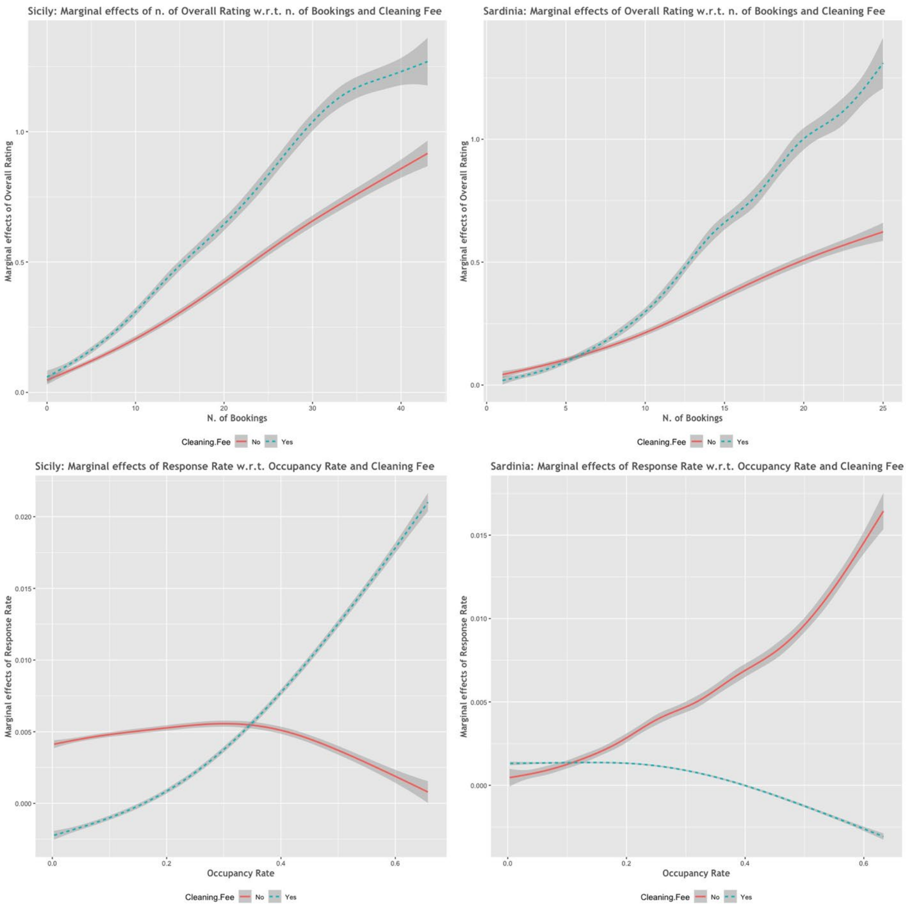

Fig. 3 Third-order interaction effects

Interesting differences between Sicily and Sardinia are notable when we focus on the joint marginal effect of the response rate, the occupancy rate and the cleaning fee on the superhost status. The bottom-left panel of Fig. 3 provides evidence that, in Sicily, increasing both the response rate and the occupancy rate causes the probability of being a superhost to increase if an extra-fee for cleaning is required, whilst there is no joint effect of the two numerical variables (or even a slightly negative effect if the occupancy rate is above 0.35) for hosts who do not require a cleaning fee. Opposite conclusions can be drawn for Sardinian hosts (bottom-right plot in Fig. 3): in this case, requiring a cleaning fee causes the probability of being a superhost to decrease, particularly if the occupancy rate is above 0.10 and/or the response rate is very low (below 0.00125). For Sardinian hosts who do not require a specific cleaning fee, instead, there is a positive joint marginal effect of occupancy rate and response rate.

In our opinion, this difference is attributable to two possible causes: the cleaning fee is implicitly included in the total fee or hosts of Sardinian rooms or apartments have highest expectations w.r.t. the quality of service and are prepared to pay more only if the quality of the service is clearly superior. Oppositely, cleaning is a discriminating factor for Sicilian hosts: for them, requiring a related extra fee and ensuring high standards for cleaning is an important discriminating factor. This finding is consistent with the idea that offering an elevate number of extra services and facilities should improve, and maintain, high quality levels in time (Liang et al. 2017).

To conclude, the empirical evidence provided in this section is consistent with Q_2. It has been demonstrated that "managerial variables", and the interactions between them and the four main variables, influence the status of superhost for hosts operating either in Sardinia or Sicily.

5 Concluding remarks

In this study, we applied the logistic and the probit models to evaluate if the findings obtained in Gunter (2018) concerning some aspects related to the Airbnb Superhost badge can be extended to other geographic areas, such as the Italian islands of Sardinia and Sicily.

These are two areas with several similarities, such as the preponderance of seaside tourism and the concentration of Airbnb accommodations in costal areas. We focus on Sardinia and Sicily because we are interested in investigating about the main features of the Airbnb phenomenon in the two largest islands of Italy. The two islands are among the most favored destinations in Italy for tourists, particularly those interested in seaside tourism, but they differ w.r.t. many cultural and economic factors that stimulate our interest in investigating about different factors influencing Airbnb's hosts.

Although in principle the two destinations might appear very similar, there are many differences to be highlighted to further motivate the present study. As an example, Sardinia's territory is characterized by a variety of ecosystems that lead to define this region "a micro-continent": they include mountains, woods, plains, largely uninhabited territories, streams, rocky coasts and long sandy beaches. Nowadays, Sardinian landscape still houses the vestiges of the Nuragic civilization.

Contrariwise, Sicily is a melting pot of a variety of different cultures and ethnicities, including the original Italic people and, among others, the Phoenicians, Carthaginians, Greeks, Romans, Byzantines, Arabs, Normans and Albanians, each contributing to the island's culture and genetic makeup. Furthermore, the Sardinian economy is today focused on the overdeveloped tertiary sector, and tourism represents the main industry of the island although its development is hindered by high costs of transportation and limited tourism planning capacity. Instead, Sicily is more easily reachable and thus it attracts more tourists from mainland Italy: although the tourist season peaks in the summer months, people visit the island all year round. Beside seasonal tourism, more Mediterranean cruise ships stop in Sicily and many wine tourists also visit the island.

The results show that the four Airbnb main variables have a real impact on the probability to become a superhost. In fact, not all the variables have the same impact: the variable Overall rating is the most important. To become a superhost, it is important to offer a high quality service, in order to be able to satisfy guests and create a nice holiday experience. Therefore, it can be argued that the model specified by Gunter (2018) is able to explain how the superhost badges are assigned, and it can be applied to different geographic areas.

Next, other variables have been included in the model to comprehend if other specific aspects can influence the probability to become superhost. Also in this case, the results evidence the importance to offer a high quality service. This aspect is certainly confirmed by the elevate impact discovered w.r.t. other variables, as for instance the requirement of additional fees for adding new guests to a previously confirmed reservation or for cleaning. Guests normally choose an Airbnb accommodation because it has less limitations compared to hotels and more traditional accommodations, and because they offer more freedom and a stronger connection with local people. Moreover, guests prefer facilities and services that make them feel like in their own homes. Hosts who want to improve their activity and become superhosts must offer a quality service to fulfill guests' needs.

Another relevant aspect is related to the capacity of the host to build a personal relationship with his guests, based on trust and respect. The two aspects can be generated only if the host takes care of his guests before and during their vacation. The creation of this relationship improves the possibility to become a superhost and it makes the real difference between an unprofessional and a professional host.

To sum up, to become a superhost it is necessary to work professionally and to offer a high quality service, support guests during the entire travel and satisfying their needs.

The results of this study have some limitations, due to the small time range considered: data were available only for 2016, and this affects the results. Future research will be addressed towards a broader time range.

Moreover, it would also be interesting to analyze more variables. This study dealt only with variables concerning aspects that can be directly controlled by hosts. As a matter of fact, though, uncontrollable variables exists that could have a remarkable impact on the probability to become a superhost. For instance, the type of accommodation cannot be directly controlled by the host, but it certainly impact the host success. Including uncontrollable variables and analyzing them together with controllable ones would certainly strengthen the model we used in this research, and would provide additional information to better define what is really important to pass to the superhost status.

Funding Funding was provided by Regione Autonoma della Sardegna (Grant No. RAS Programmazione Unitaria 2007/2013 - P.O. FESR 2007/2013, Interventi a sostegno della competitività e dell'innovazione).

References

Aznar, J.P., Sayeras, J.M., Rocafort, A., Galiana, J.: The irruption of Airbnb and its effects on hotel profitability: an analysis of Barcelona's hotel sector. Intang Cap **13**(1), 228–237 (2017)

Blal, I., Singal, M., Templin, J.: Airbnb's effect on hotel sales growth. Int. J. Hosp. Manag. **73**, 85–92 (2018)

Boros, L., Dudás, G., Kovalcsik, T., Papp, S., Vida, G., et al.: Airbnb in Budapest: analysing spatial patterns and room rates of hotels and peer-to-peer accommodations. GeoJ. Tour. Geosites **21**(1), 26–38 (2018)

Bridges, J., Vásquez, C.: If nearly all Airbnb reviews are positive, does that make them meaningless? Curr. Issues Tour. **21**(18), 2057–2075 (2018)

Choi, K.H., Jung, J.H., Ryu, S.Y., Do Kim, S., Yoon, S.M.: The relationship between Airbnb and the hotel revenue: in the case of Korea. Indian J. Sci. Technol. **8**(26), 1–8 (2015)

Dudás, G., Boros, L., Kovalcsik, T., Kovalcsik, B.: The visualisation of the spatiality of Airbnb in Budapest using 3-band raster representation. Geogr. Tech. **12**(1), 23–30 (2017)

Dudás, G., Vida, G., Kovalcsik, T., Boros, L.: A socio-economic analysis of Airbnb in New York City. Reg. Stat. **7**(1), 135–151 (2017)

Ert, E., Fleischer, A., Magen, N.: Trust and reputation in the sharing economy: the role of personal photos in Airbnb. Tour. Manag. **55**, 62–73 (2016)

Fahrmeir, L., Tutz, G.: Multivariate Statistical Modelling Based on Generalized Linear Models, Springer Series in Statistics. Springer, New York (1994)

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., Goodwill, A.: Pricing in the sharing economy: a hedonic pricing model applied to airbnb listings. J. Travel. Tour. Mark. **35**(1), 46–56 (2018)
Gunter, U.: What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area. Tour. Manag. **66**, 26–37 (2018)
Gurran, N., Phibbs, P.: When tourists move in: how should urban planners respond to Airbnb? J. Am. Plann. Assoc. **83**(1), 80–92 (2017)
Gutierrez, J., Garcia-Palomares, J.C., Romanillos, G., Salas-Olmedo, M.H.: Airbnb in tourist cities: comparing spatial patterns of hotels and peer-to-peer accommodation. arXiv preprint: arXiv:1606.07138 (2016)
Gutiérrez, J., García-Palomares, J.C., Romanillos, G., Salas-Olmedo, M.H.: The eruption of Airbnb in tourist cities: Comparing spatial patterns of hotels and peer-to-peer accommodation in Barcelona. Tour. Manag. **62**, 278–291 (2017)
Gutt, D., Herrmann, P.: Sharing means caring? Hosts' price reaction to rating visibility. In: ECIS 2015 Research-In-Progress Papers, vol. 54, pp. 1–13 (2015)
Guttentag, D.: Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. Curr. Issues Tour. **18**(12), 1192–1217 (2015)
Guttentag, D.A., Smith, S.L.: Assessing Airbnb as a disruptive innovation relative to hotels: substitution and comparative performance expectations. Int. J. Hosp. Manag. **64**, 1–10 (2017)
Leeper, T.J.: Interpreting regression results using average marginal effects with R's margins. Techical Report. https://cran.r-project.org/web/packages/margins/index.html (2017). Accessed 05 Aug 2019
Li, J., Moreno, A., Zhang, D.J.: Agent behavior in the sharing economy: evidence from Airbnb. Ross School of Business Working Paper Series (1298), pp. 1–33 (2015)
Liang, S., Schuckert, M., Law, R., Chen, C.C.: Be a "Superhost": the importance of badge systems for peer-to-peer rental accommodations. Tour. Manag. **60**, 454–465 (2017)
Lutz, C., Newlands, G.: Consumer segmentation within the sharing economy: the case of Airbnb. J. Bus. Res. **88**, 187–196 (2018)
Ma, X., Hancock, J.T., Mingjie, K.L., Naaman, M.: Self-disclosure and perceived trustworthiness of Airbnb host profiles. In: CSCW '17 Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 2397–2409 (2017)
Martin-Fuentes, E., Fernandez, C., Mateu, C., Marine-Roig, E.: Modelling a grading scheme for peer-to-peer accommodation: stars for Airbnb. Int. J. Hosp. Manag. **69**, 75–83 (2018)
McNamara, B.: Airbnb: A not-so-safe resting place. J. on Telecomm. & High Tech. L. **13**, 149–170 (2015)
Oskam, J., Boswijk, A.: Airbnb: the future of networked hospitality businesses. J. Tour. Futures **2**(1), 22–42 (2016)
Powers, D., Xie, Y.: Statistical Methods for Categorical Data Analysis. Emerald Group Publishing, Bingley (2008)
Quattrone, G., Proserpio, D., Quercia, D., Capra, L., Musolesi, M.: Who benefits from the sharing economy of Airbnb? In: Proceedings of the 25th international Conference on World Wide Web, pp. 1385–1394. International World Wide Web Conferences Steering Committee (2016)
Roelofsen, M., Minca, C.: The Superhost. Biopolitics, home and community in the Airbnb dream-world of global hospitality. Geoforum **91**, 170–181 (2018)
Wang, D., Nicolau, J.L.: Price determinants of sharing economy based accommodation rental: a study of listings from 33 cities on Airbnb.com. Int. J. Hosp. Manag. **62**, 120–131 (2017)
Zervas, G., Proserpio, D., Byers, J.: A first look at online reputation on Airbnb, where every stay is above average. Available at SSRN (2015). https://doi.org/10.2139/ssrn.2554500
Zervas, G., Proserpio, D., Byers, J.W.: The rise of the sharing economy: estimating the impact of Airbnb on the hotel industry. J. Mark. Res. **54**(5), 687–705 (2017)
Zhang, L., Yan, Q., Zhang, L.: A computational framework for understanding antecedents of guests' perceived trust towards hosts on Airbnb. Decis. Support Syst. **115**, 105–116 (2018)