ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Master Thesis Marketing

"To Airbnb or not to Airbnb: What Makes one Place More Bookable than Another?"

Abstract

The research tries to bridge two gaps in prior literature using the signaling theory. Firstly, the effect of contractual strictness on the performance of sharing rental accommodations. Secondly, the effect of quality indicators on the performance distinguishing between host- and listing specific quality indicators. Therefore, this research investigates what the effect of contractual strictness, proxied by cancellation policy strictness, and the quality indicators, proxied by Open Badges, is on the performance of listings, proxied by the estimated occupancy rate, on sharing rental accommodation platforms such as Airbnb.com. Additionally, because prior literature shows evidence of discrimination, the research also investigates if the gender of a host moderates these effects. The research question is answered by testing two theory-based hypotheses using a regression model, which is tested separately for the cities Amsterdam, Paris and London. It was found that the contractual strictness has a positive effect on the performance to a certain extent. Regarding the Open Badges, solely the superhost badge has a strong positive effect on the performance. These effects were not moderated, with the exception of London, by the gender of the host. However, the inconsistency in significant results between different cities suggests that further research is needed.

Name student: Halima Vanhee
Student ID number: 428757

Supervisor: Dr. (Nuno) N.M. Almeida Camacho
Second assessor: Prof. Dr. (Vardit) V. Landsman - Schwartz

July 15th, 2020

Table of Contents

1. Introduction

"*As people's access to the Internet grows we're seeing the sharing economy boom – I think our obsession with ownership is at a tipping point and the sharing economy is part of the antidote for that*." – Richard Branson, founder of Virgin Group (Lawrie, 2016).

The rise of the Internet and digital technology carried many advantages. It enabled: the diffusion of richer and participative information, transparency, diversity, competitive pricing and cost reduction (e.g. search costs) (Varadarajan and Yadav, 2002; Schor, 2014; Tiago and Verissimo, 2014). Consequentially, it has given rise to the sharing economy, which is an economic model defined as peer-to-peer marketplaces from which certain grew quickly and became well-known brands worldwide such as Airbnb and Uber (Fraiberger and Sundararajan, 2015). In the hospitality industry, Airbnb is the largest peer-to-peer platform and hotels represent the incumbents (Farronato and Fradkin, 2018).

Moreover, Airbnb was founded in 2008 and has grown to a marketplace with more rooms than any hotel group in the world. Platforms such as Airbnb have expanded the hospitality market and increased demand (Li and Srinivasan, 2019). Users benefit from platforms such as Airbnb thanks to the offering of differentiated products in comparison with hotels. In addition, the supply of rooms is expanding over time as it is directly competing with hotels (Farronato and Fradkin, 2018). Furthermore, the presence of peer-to-peer platforms for rental accommodations such as Airbnb lowers the revenues of hotels (Zervas et al., 2017). The main explanation is that the peer supply elasticity is twice as high as the hotels' supply elasticity on average, which means that the consumer-supply is more responsive to demand (Zervas et al., 2017; Farronato and Fradkin, 2018).

The arrival of this particular marketplace has called attention and discussion regarding its ambiguous definition and varied regulation (i.e. regulation over the world), mainly due to the novelty of the concept (Rauch and Schneider, 2015; Farronato and Fradkin, 2018; Frenken and Schor, 2019). The concern stems from observing the sharing economies' growth and threat in becoming monopolies as they are characterized by network externalities (i.e. change in value of product or service with an increasing number of users), which enables higher profit margins and threatens competition (Katz and Shapiro, 1985; Dellaert, 2019; Frenken and Schor, 2019). In this paper the definition for "sharing economy" is adopted from Frenken et al. (2015): "*Consumers granting each other temporary access to under-utilized physical assets (i.e. "idle capacity"), possibly for money*."

Furthermore, the sharing economy has three defining characteristics: consumer-to-consumer interaction (i.e. C2C), temporary access (i.e. temporary stay) and physical (shareable) goods (see Appendix A, Figure A.1) (Frenken and Schor, 2019). Examples of these "shareable goods" are cars and homes (Benkler, 2004). From the three characteristics one is most obvious: consumer-to-consumer interaction (i.e. C2C) (see Appendix A, Figure A.1) (Frenken and Schor, 2019). A consumer acts as host (i.e. supplier) and can easily offer his/her spare space (i.e. idle capacity) to book for a determined time (i.e. temporary) to another consumer active as a prospective guest.

The Internet has not only decreased the transaction costs, search costs, but also contract costs involved in this "stranger sharing" (Schor, 2014). This consequentially made sharing among peers easier and more efficient (Benkler, 2004; Schor, 2014; Frenken and Schor, 2019). The "stranger sharing" is facilitated by platforms such as Airbnb that adopted standardized contracts and online payment services, which regularizes the transactions (Schor, 2014). In addition, most sharing economy platforms such as Airbnb show past behavior of users (i.e. through reviews and status symbols), which promotes trustworthiness. In doing so, this additionally lowers transaction costs and lowers risk for both parties (Frenken and Schor, 2019).

Therefore, the host can disclose what he/she offers (i.e. property characteristics) and is free to determine the house rules (see Appendix A, Figure A.2). These house rules influence the relationship between guest and host. Additionally, the guests use these rules to evaluate the host and his/her listing (Fornell et al., 1996; Vos et al., 1998; Zeithaml, 2000; Jiang and Rosenbloom, 2005). Therefore, it is important to understand both the host (supply) and guest (demand) factors that determine transactions in these growing sharing economy markets, such as Airbnb. These house rules can be considered as contractual clauses, which are specific sections within a contract (e.g. refund/cancellation policy, check-in/checkout time frame and deposit fee) (McMeel, 2007). Besides the standardized contracts and payment service, Airbnb also aids hosts "set expectations" of the guests by obliging/enforcing guests to agree to the hosts' rules (i.e. contractual clauses) before booking (i.e. if rule/clause is broken, the host can cancel the reservation) (Airbnb, n.d.).

Moreover, there is a lack of research that looks into the effect of host strictness in the sharing economy. Reasons for why a host could be more inclined to stricter house rules can be because he/she values capital preservation over potential higher returns (e.g. host is risk-averse) or because the host is not flexible with his/her time (e.g. host is busy or host preferences) (Skaperdas, 1991). However, previous research regarding warranties and refunds found that for services such as airlines and hotels it is beneficial to offer partial refunds (i.e. flexible refund). This encourages buyers to stop looking for alternatives and promotes trust. Additionally, this refund increases the efficiency and utilization but also intensifies competition (Srivastava and Lurie, 2001; Xie and Gerstner, 2007; Guo, 2009).

The advancements in digital technology have affected and created drivers of consumer decision making (Schor, 2014, Xie et al., 2014; Zervas et al., 2015). Additionally, it affected service perception and evaluation. Research showed that it is more beneficial to offer a personal relationship-based service as opposed to a quality-based service to maintain quality at customer expectations (Tiago and Verissimo, 2014; Pansari and Kumar, 2016; Huang and Dev, 2019).

Digital advancements enabled the creation of Open Badges, which are "a validated indicator of accomplishment, skill competency and quality" and a common characteristic of the sharing economy (Jovanovic and Devedzic, 2015). These Open Badges are measurable and verifiable for everyone and therefore associated with transparency, decrease uncertainty and promote trust. The prospective guest can rely on these badges as cues for quality when choosing a listing (see Appendix B, Figure B.1 – B.5). The Open Badges are also beneficial to hosts to credibly convey and so attract prospective guests. The host can apply for these, but only earn them when he/she meets the requirements/credentials.

The sharing economy is creating enormous amounts of wealth, carries a social, trusted and feel-good image. However, it has also increased peer-to-peer discrimination (Bertrand and Mullainathan, 2004; Edelman and Luca, 2014; Ge et al., 2016; Cansoy and Schor, 2017; Edelman et al., 2017). Research showed exclusionary behavior in the choice of a trading partner on peer-to-peer platforms hurting ratings, reviews and the prices charged towards both guests and hosts of color (Cansoy and Schor, 2017; Schor et al., 2016). However, there is a lack of research regarding gender bias (i.e. inclination or discrimination towards a particular gender) within a sharing economy context. Previous research found that gender bias in the context of education and workplace is mainly due to a gender imbalance and gender-stereotypical associations that set expectations (Bennett, 1982; Eagley and Karau, 2002; Boring, 2017). There has been no research that investigated the various listing profile attributes (i.e. what the guests see before booking) in interaction with the gender of the host. Are listings with strict rules more tolerated (i.e. booked) if the host is female?

Prior literature faces two important gaps on both the supply side and the demand side. Research from the suppliers' perspective is lacking on a deeper level, especially regarding the contractual clauses in a sharing economy context (see Appendix C, *Table* C.1). It is not yet known what the effect of host inclined to impose stricter rules (e.g. host is risk-averse or has a lack of time) is on the performance of his/her listing within a sharing economy context. Nonetheless, there is much research regarding the relationship between the Open Badges and the performance of companies in a sharing economy context. Even though more attention is given to personalized customer relationships in previous research, there has not been any research determining what is more valued on a sharing rental accommodation platform such as Airbnb: the excellence of the host or his/her listing? A comparison between those two is missing. Lastly, much research discloses exclusionary behavior in the sharing economy, but research investigating the relationship between the various components of the profile separately in interaction with the gender of the host has not been done before.

In this thesis, I try to contribute to the literature by offering an early attempt to bridge these two gaps. Specifically, I propose a typology based on the attributes the host can implement/expand in the listing profile and what the prospective guest can see. Next, I develop theory-based hypotheses tested with secondary data scraped from Airbnb.com, which is publicly available on Inside Airbnb (http://insideairbnb.com). The research uses the estimated occupancy rate as a performance measure of rental accommodation in the sharing economy (see Appendix D, Table D.1). This percentage shows the ratio of the estimated occupied nights of a listing (i.e. booked) to the year (i.e. 365 days). Furthermore, the investigation will look at the strictness of the listing by focusing on the various cancellation policies. Additionally, I investigate whether the gender of the host moderates these effects or not. Besides the strictness, I attempt to reveal what quality-signaling attributes of a listing profile (i.e. Open Badges) are more important for the performance by distinguishing the host-specific from the list-specific ones. Additionally, I assess whether gender moderates this effect as well.

To develop the theory-based hypotheses, I rely on the signaling theory. This theory rides on the premise that an agent decides how to credibly convey information about him/herself to the principal, who decides how to interpret the signal (Ross, 1973; Spence, 1978; Connelly et al., 2011).

This leads to the research question:

"*What influence do strictness and quality signals have on the performance of sharing rental accommodations and does it matter whether or not these are coming from a male or female host? – an Airbnb Case*"

Two Hypotheses are tested in order to answer the research question. The research paper is structured in a way that the theoretical framework follows this Introduction (i.e. chapter 2). After the Theoretical Framework, the Data and Methodology are laid out (i.e. chapter 3). Lastly, the findings are shown and discussed in the sections: Results (i.e. chapter 4) and Discussion (i.e. chapter 5).

2. Theoretical Framework

To answer the research question, theory-based hypotheses are built upon the signaling theory. The theory predicts that the more (credible) information there is on a listing profile, the more it promotes trust and so attracts guests ((Ross, 1973; Spence, 1978; Connelly et al., 2011). An important advance on signaling theory was George Akerlof's "market for lemons" theory, which is built on the premise that truthful, credible disclosure (i.e. credibly signaling information) is too expensive (Akerlof, 1978). However, digital technology advancements improved software, leading to easier and less expensive forms of disclosure and signaling (Akerlof, 1978; Lewis, 2011). In the context of sharing economy, two such forms are apparent and the focus of my thesis: the *contractual clauses* designed by sellers, which I argue can function as a signaling device and *Open Badges* often available in sharing economy platforms, I discuss each in turn.

2.1. The Signaling Effects of Contractual Strictness on Sharing Economy Platforms

Contractual clauses are entirely determined by the host and disclose house rules to prospective guests that need to accept these (McMeel, 2007). There are hosts that value the preservation of capital more than others (e.g. they are more risk-averse) or are restricted by time (e.g. host has a busy job or host preferences). Therefore, it can be expected that these hosts will be relatively more inclined to impose stricter house rules (Skaperdas, 1991). Nevertheless, in doing so the host might signal distrust towards prospective guests causing a negative signaling effect of strict contractual clauses (i.e. contractual strictness). Previous research regarding warranties and refunds has found that flexible refunds promote trust and discourage further search for better alternatives (Srivastava and Lurie, 2001; Xie and Gerstner, 2007; Guo, 2009).

Moreover, flexible refund schemes have been referred to as "double-edged swords" since they increase efficiency of utilization, but also intensify competition (Guo, 2009). The competition on Airbnb can be expected to be intense as the centralized sharing accommodation platform has low barriers of entry and a high threat of substitutes (Porter, 2008). Research implies that refund schemes can intensify this, which would consequentially hurt the performance.

Nonetheless, stricter clauses can also have a positive signaling effect. Stricter clauses can be translated in for example a narrow check-in/checkout timeframe, strict cancellation policy or, a plethora of (relatively high) fees (e.g. service fee, deposit security). Additionally, price is considered as a contractual clause in this research since it is entirely determined by the host (i.e. no guarantee of accurate pricing). An additional reason is that the dataset is merely a "snapshot", therefore price changes are not observable. The contractual strictness affects the *perceived quality* and should not be confused with *objective quality* (e.g. reflected by Open Badges) (Zeithaml, 1988). Objective quality refers to the actual technical superiority and merit, which is measurable and verifiable with predetermined standards. Perceived quality on the other hand refers to subjective interpretation and differs per person (Hjorth-Anderson, 1984; Monroe and Krishnan, 1985; Curry and Faulds, 1986; Zeithaml, 1988). The Open Badges are examples of objective quality indicators as these are measurable and verifiable.

Previous research showed that price and brand are indeed performance attributes, but they rather reflect prestige/exclusivity (i.e. perceived quality) than quality (i.e. objective quality) (Zeithaml, 1988; Brucks et al., 2000; Zeithaml, 2000). This implies that a listing that is perceived as relatively expensive (i.e. price, switching costs caused by strict cancellation policy, various fees) will be seen as prestigious and exclusive (i.e. not everyone can afford it). This is in contrast to Open Badges that serve as verifiable and measurable indicators referring to the superiority in terms of excellence (e.g. location, cleanliness).

Therefore, attributes such as contractual strictness could be signaling two effects: either prestige/exclusivity and so attract guests or distrust and repel guests (Ross, 1973; Spence, 1978; Zeithaml, 1988; Weiner, 2000; Srivastava and Lurie, 2001; Guo, 2009; Connelly et al., 2011; Oghazi et al., 2018). For this research, the contractual strictness is operationalized by the observed strictness of a listing, for which the cancellation policy strictness serves as a proxy.

2.2. The Signaling Effects of Open Badges on Sharing Economy Platforms

As discussed above digital technology advancements – some of which particularly visible in sharing economy platforms – led to the emergence of easier and cheaper forms of disclosure and signaling (Akerlof, 1978; Lewis, 2011). The Airbnb platform frequently gathers massive amounts of data, which makes it easy to obtain the data needed to classify both hosts (e.g. response time, reviews, etc.) and guests (reservations canceled, reviews by hosts, etc.). Many of these data points also ensure that hosts, for instance, truly meet the quality requirements and promises they make.

In such a context, an Airbnb host needs to convey guests that the listing and his/her competences are as great as he/she signals. Airbnb solves potential uncertainty by offering several quality-status symbols, Open Badges, for which hosts can apply for and publish when the requirements are met (see Appendix B, Figures B.1, B.3 and B.5). Some Open Badges are standard on every profile such as cleanliness and communication score (see Appendix B, Figures B.4 and B.5). The requirements are verified by recorded data and algorithms, which cannot be manipulated by the host and easily verifiable for both host and guest. These badges are beneficial for both parties: it allows the suppliers to differentiate themselves from the competition and decreases uncertainty and risk for the demand (Akerlof 1978; Lewis, 2011).

However, there are two types of Open Badges that a host can obtain: *host-related* and *listing-related* badges. Examples of host-related Open Badges are Superhost, Communication, Accuracy and Check-in (see Appendix B, Figures B.1, B.4 and B.5). Examples of listing-related Open Badges are Rare find, Cleanliness, Location and Value (see Appendix B, Figures B.1, B.4 and B.5). The "Plus" Open Badge is a combination of host and listing competencies and is only available in predetermined locations (see Appendix B, Figure B.3).

Given that obtaining these badges still involves some costs for hosts (e.g. time to receive guests well and ensure an excellent experience, cleaning costs, investing in the property, etc.), this leads to a critical question for hosts – out of these two types of Open Badges, which one is more of a deal-breaker to guests booking: how great listing: how great the listing is or how great the host is?

2.3. Discrimination in Signaling on Sharing Economy Platforms

Moreover, there is also evidence of potential discrimination against the host on the sharing economy platforms, which may influence the effectiveness of signaling efforts by hosts. For instance, existing literature shows that there is exclusionary behavior in the choice of a trading partner on peer-to-peer platforms hurting both guests and hosts of color (Edelman and Luca, 2014; Schor et al., 2016; Cansoy and Schor, 2017; Edelman et al., 2017).

There is extensive research regarding gender bias in several areas, such as the workplace and in education. Research in education showed gender bias of students in teacher evaluations. Additionally, it showed that students were biased towards their own gender (Ferber and Huber, 1975; Boring 2017). This can be explained by the fact that students evaluate and rate teachers in different teaching dimensions and expectations attached to particular gender stereotypes (Bennett, 1982; Boring, 2017). Regarding leadership roles, gender bias occurs because it is perceived as incompatible with female gender roles. Since leadership roles are mostly male-dominated it sets certain associations and so expectations, which consequentially leads to less positive attitudes towards female leaders (Eagley and Karau, 2002).

Therefore, previous research explains that gender bias results from an imbalance of gender, but also perceived incongruities between gender and functional roles. This means that people have gender-stereotypical expectations that affect their decision-making. The reason why is because gender is a characteristic that provides the strongest and easiest base to classify other people (Fiske et al., 1991; Stangor et al., 1992; van Knippenberg et al., 1994; Eagley and Karau, 2002). This process happens automatically and is even stronger than age, race and employment (Banaji et al., 1993; Banaji and Hardin, 1996; Eagley and Karau, 2002).

Hence, gender bias is mostly generated through gender imbalance/dominance and gender-stereotypical characteristics that set expectations. Female Airbnb hosts represent about 56% and are on average higher-earning hosts as opposed to male hosts (Airbnb, 2019). Interestingly, senior women are consistently the best rated hosts relative to the others (iPropertyManagement, n.d.). The exclusionary behavior has not yet been investigated in interaction with the separate components of a listing.

Therefore, this research will study whether there is gender discrimination against the hosts and how this affects the relation of contractual strictness and Open Badges with a listing's performance (see Figure 1).

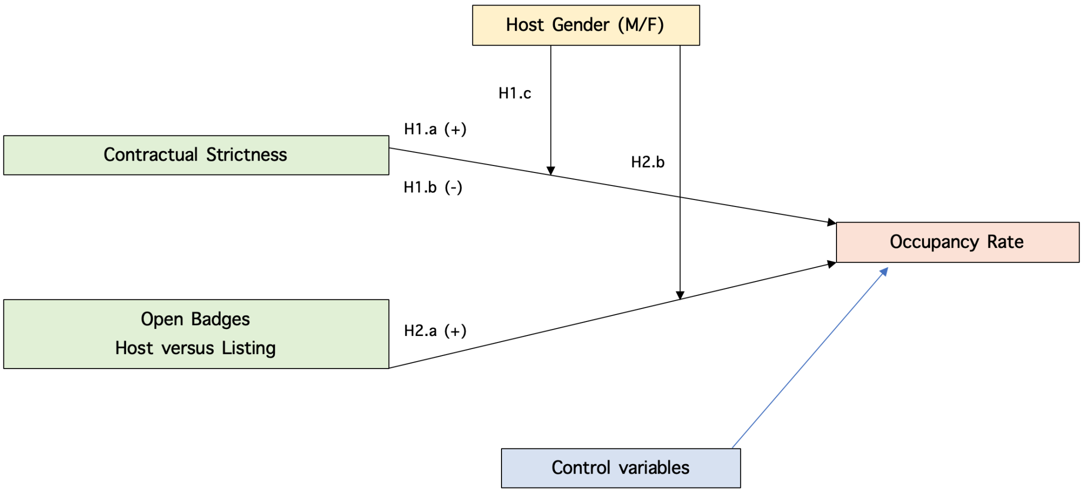


Figure 1. Conceptual Framework.

2.3.1. Discrimination and Contractual Strictness

Firstly, regarding the strictness of the contractual clauses, the host gender can intensify certain effects on the booking decision of the guest (see Figure 1). As seen from previous research there are certain stereotypes attached to gender that set expectations. Characteristics that are associated with the female role are: warm, nurturing and empathetic (Bennett, 1982). However, if women in certain domains do not meet these expectations, they get punished for it (Bennett, 1982; Boring, 2017). A related topic to this is in leadership style, the so-called "double bind dilemma", which states that female leaders are either seen as competent or likeable, but not both (Catalyst, 2007; Achinstein and Ogawa, 2012).

Therefore, this implies that on the one hand female hosts imposing stricter clauses could be perceived as deviating from the nurturing and warm gender-stereotypical characteristics and so be punished for it in the evaluations when these expectations are not met (i.e. negative effect). On the other hand, there is not really a gender imbalance in Airbnb hosts (i.e. 56% female hosts) and the sharing economy is rather new (i.e. stereotypical characteristics are maybe not established yet, no established gender bias yet). Additionally, there might even be understanding or sympathy in favor of women because of recent events such as the "#METOO" movement (i.e. positive effect) (Mendes et al., 2018).

2.3.2. Discrimination and Open Badges

Further, discrimination in relation to Open Badges is interesting to evaluate since these validated quality indicators were initially designed to promote trust and set expectations with prospective guests. As aforementioned, gender also affects certain expectations based on stereotypical characteristics associated with gender (see Figure 1).

Anne Boring (2017) shows gender bias in the context of teaching evaluation. Students value the more time-consuming teaching dimension for female teachers more (e.g. organization course content) (Boring, 2017). Additionally, female teachers were rated higher when certain expectations caused by gender-stereotypical characteristics were met, such as warmth, charismatic and have interpersonal skills (Bennett, 1982). This is opposed to male teachers, that are evaluated more favorably in less time-consuming teaching dimension for male teachers more (e.g. quality of animation, class leadership skills) (Boring, 2017).

Additionally, research found in the context of management and leadership roles that women are subject to higher standards of competence due to male job domination and so the perceived deviation of stereotypes. This implies that in this particular context, women need to prove their abilities more as opposed to men (Catalyst, 2007; Elsesser and Lever, 2011). The women that break the stereotype, are perceived as "cold" or as "trying too hard" (Catalyst, 2007; Fine, 2010). However, as seen before the Airbnb platform does not have a gender imbalance (i.e. 56% female hosts) and it is relatively new (i.e. stereotypical characteristics are maybe not established yet).

Further, the effect of the host gender could intensify the effect of Open Badges because there might be certain stereotypical characteristics attached to the host gender like in the teacher's evaluation context (Boring, 2017) (see Figure 1). For example, the guest could expect more organizational and personal skills (e.g. check-in experience, communication and cleanliness) from female hosts as opposed to male hosts. The Open Badges serve as validated proof of host and listing competences, but there is a lack of research that investigates the interaction of Open Badges with gender.

2.4. Theoretical Framework - Wrap Up

Having laid out the theoretical arguments in the previous sections, I directly synthesize the theoretical expectations in a clear set of hypotheses with the aim of answering the research question below (see Figure 1). For a detailed overview of the theorized mechanism, please refer back to the theory sections (Theoretical framework *2.1 - 2.3.2*).

"*What influence do strictness and quality signals have on the performance of sharing rental accommodations and does it matter whether or not these are coming from a male or female host? – an Airbnb Case*"

Hypothesis 1:

H1.a. *A strict cancellation policy has a positive statistically significant effect on the occupancy rate.*

H1.b. *A strict cancellation policy has a negative statistically significant effect on the occupancy rate.*

H1.c. *The relationship between strict cancellation policy and the occupancy rate will be stronger among women.*

Hypothesis 2:

H2.a. *The Open Badges have a positive statistically significant effect on the occupancy rate.*

H2.b. *The relationship between Open Badges and the occupancy rate will be stronger among women.*

3. Data and Methodology

3.1. Data Collection

In order to test the theoretical predictions, data is collected through two sources. Firstly, the data about the Airbnb listings and its properties is needed. Therefore, the main datasets (i.e. "listings.csv.gz" file per city: $N_{Amsterdam} = 19{,}278$, $N_{Paris} = 59{,}320$ and $N_{London} = 86{,}358$) are retrieved from Inside Airbnb (http://insideairbnb.com). The data of the listings were scraped on the 9th of May 2020. Inside Airbnb frequently scrapes the Airbnb website and provides data for multiple cities online. The datasets cover almost all properties of a listing that is shown and described on an Airbnb listing profile. A dataset for each city (i.e. Amsterdam, Paris and London) is used separately for this research instead of merged into one dataset since each city has its own short-term rental laws and requirements. For example, the maximum length of stay in order for a listing to be categorized as a short-term rental for each city is different (see Appendix D, Table D.2).

3.1.1. Dependent Variable: Occupancy Rate

The dependent variable in the research serving as a proxy for performance is the occupancy rate. This estimate represents how many days of the year the listing is booked. This variable is visualized per listing on the Inside Airbnb website (http://insideairbnb.com) but is not provided in the public datasets (i.e. "listings.csv.gz" files). However, the website provides the occupancy model which Inside Airbnb uses in the disclaimers (i.e. "San Francisco Model"), which is replicated for this research (see Appendix D, Table D.1).

Nonetheless, this represents a limitation to take into account since the model does not include some important variables such as availability (i.e. how often the host puts the listing available), which could be used to normalize the occupancy rate. For example, it is unfair to compare the "days occupied per year" of a listing that is put available for three weeks a year versus a listing put available for two months a year with respect to performance. It would be more accurate if the data showed the history of the availability and booked days to make a comparison amongst listings fairer.

Fortunately, the variable *estimated availability* is provided, yet, this is the availability that the host has filled in at the time it was scraped (i.e. 9th May 2020) and can still be changed later in the year by the host (i.e. extend or shorten availability). Therefore, I use this variable, i.e. *estimated availability*, to create a control variable *has_high_availability*, to control whether or not these listings have an availability that respects the short-term rental period (see Appendix D, Table D.2). Those that exceed this period are required to take extra administrative measures (i.e. costs) in order to legally rent out on Airbnb (e.g. register or get a license). These highly available listings are assumed to be more professional as it indicates that the listing is more of an additional income than solely sharing empty spaces.

3.1.2. Moderator: Host Gender

Moreover, since these datasets include the name of the host for each listing, the gender of each host can be determined. For this, an additional dataset was created with names and the corresponding gender. This dataset is created through merging already existing datasets (e.g. https://data.world/howarder/gender-by-name). The names dataset consists of 113,687 names ($N_{FemaleNames} =$ 69,738 names; $N_{MaleNames} =$ 43,949 names).

In the dataset containing the Airbnb listings, I use the created names dataset to identify the gender of the host while excluding unisex names (i.e. for both male and female) and couples. After matching the names with the gender, excluding duplicates and outliers, the final Airbnb listing sample sizes were the following for each city: $N_{Amsterdam} = 3{,}909$, $N_{Paris} = 16{,}888$ and $N_{London} = 25{,}873$.

3.2. Measurement

3.2.1. Dependent Variable: Occupancy Rate

As aforementioned, the dependent variable *occupancy_rate* is the proxy for the performance of the sharing rental accommodations. This is estimated according to the occupancy model (i.e. “San Francisco Model”) from Inside Airbnb (see Appendix D, Table D.1).

3.2.2. Independent Variable: Cancellation Policy Strictness

Furthermore, the explanatory variable for the first hypothesis is the categorical variable *cancellation policy strictness*, which is the proxy for contractual strictness. This variable is broken down in three types: *flexible*, *moderate* and *strict policy*. The explanation of each type of cancellation policy is shown below in Table 1. The three types are treated as separate dummy variables to observe an effect of each separately with the *flexible policy* variable as a reference (i.e. not included). The variable *strict policy* contains three types of strict policies. This is because the sample size of listings having a super strict policy (i.e. *super_strict_30/60*) is relatively small, and they are also strict. Therefore, these are grouped together under the variable *strict policy*.

Table 1. The Types of Cancellation Policies Offered on Airbnb.

Cancellation Policies	
Flexible policy	Free cancellation for 48 hours. After that, cancel up to 24 hours before check-in and get a full refund, minus the service fee.
Moderate policy	Free cancellation for 48 hours. After that, cancel up to 5 days before check-in and get a full refund, minus the service fee.
Strict policy	Includes the three strict policies: - *strict_14_with_grace_period* i.e. free cancellation for 48 hours. After that, cancel up to 7 days before check-in and get a 50% refund, minus the service fee. - *super_strict_30* i.e. cancel up to 30 days before check-in and get a 50% refund, minus the service fee. - *super_strict_60* i.e. cancel up to 60 days before check-in and get a 50% refund, minus the service fee.

3.2.3. Independent Variable: Open Badges

Regarding the second hypothesis, the Open Badges for both host and listing are the explanatory variables. The Open Badges and their descriptions can be found in Table 2 below. Besides the dummy variable indicating whether or not the host is a *superhost*, the score badges are discrete values (e.g. *score_location*).

However, on the official Airbnb website, each listing profile has these scores on a scale from 1-5 (i.e. continuous) and not 2-10 (i.e. discrete). Because of an unknown reason Inside Airbnb has doubled and rounded off these ratings. Hence, this represents another limitation to take into account since it takes away variation and distinction among listings. For example, two listings with an Open Badge score of cleanliness of 4.3/5 and 4.7/5, will both have a score of 9/10 in the Inside Airbnb dataset.

Table 2. The Host- and Listing-Related Open Badges.

Host-related Open Badges		**Listing-related Open Badges**	
Host is superhost	He/she is a great host that meets who requirements according to *Figure* 3 in the appendix	Score cleanliness	Discrete value between 2 and 10. Guests' opinion regarding the cleanliness of the listing.
Score communication	Discrete value between 2 and 10. Guests' opinion regarding the host' communication.	Score value	Discrete value between 2 and 10. Guests' opinion regarding the value provided by the listing in comparison to the price.
Score check-in	Discrete value between 2 and 10. Guests' opinion regarding the check-in experience with the host.	Score accuracy	Discrete value between 2 and 10. Guests' opinion regarding the accuracy of the listings' presentation in comparison to his/her experience.
		Score location	Discrete value between 2 and 10. Guests' opinion regarding the location of the listing.

3.2.4. Moderator: Host Gender

To assess the influence of cancellation policy strictness and Open Badges on the performance of sharing rental accommodations and whether or not the gender of the host amplifies these effects, I use a regression model including main effects and interactions with the host's gender. The gender of the host is identified based on the provided name of the host, which is matched to the associated gender. The binary variable *host_gender* takes the value of 1 indicating the host is female. To fully interpret the moderation of the variable *host_gender* with the score badges (e.g. *score_checkin*) represented by interaction terms, I test the effect of the variable *host_gender* on the variable *occupancy_rate* at various levels of these Open Badge scores (i.e. discrete values ranging from 2 to 10) using margin plots (see Appendix E, Figures E.1 – E.5). The margin plots show how the coefficient of the variable *host_gender* (on the y-axis, with 95% confidence interval) varies for different values of each Open Badge score variable (on the x-axis), and for which of these score levels the effect is significant (indicated in purple on the margin plots).

3.2.5. Control Variables

Furthermore, several control variables were added to the model, which can be seen in Table 3 below.

Table 3. List of Control Variables Used in the Hypotheses.

	Control variables	**Explanation**
1	'log_price_per_person'	The (logarithmic) price ('price') divided by capacity ('accommodates') of the listing.
2	'log_security_deposit'	A (logarithmic) sum to cover damages to the property that occurred during their stay.
3	'log_cleaning_fee'	(Logarithmic) Fee to cover the costs of cleaning the listing after guest departs.
4	'log_extra_people_fee'	(Logarithmic) Fee to cover additional person exceeding the accommodation capacity.
5	'log_distance_cost'	(Logarithmic) distance from the central zone (see appendix, *Figure* 9.1 – 9.3).
6	'instant_bookable'	Dummy variable, 1= no approval by host process or 0= approval by host need.
7	'host_is_local'	Dummy variable, 1= the location of the host is in the city (Amsterdam, London and Paris) or 0= host lives outside the city/country.
8	'host_multiple_listings'	Dummy variable, 1= number of listings owned by host exceeds 1or 0= host owns 1 listing (< 2 listings).
9	'has_high_availability'	Dummy variable, 1= availability of the listing exceeds the cities' short-term rentals threshold (e.g. Amsterdam 60 days, London 90 days and Paris 120 days) or 0= below or equal to cities' threshold.
10	'accommodate'	The capacity of each listing.
11	'neighborhood_n'	Categorical value, each neigbourhood of the city.
12	'property_type_n'	Categorical value, e.g. house, boat, bed and breakfast.
13	'room_type_n'	Categorical variable, e.g. entire house/apartment, private room, shared room and hotel room.

Most variables were already included except for *distance_cost* (i.e. distance from center), *price_per_person* (i.e. price divided over the capacity of the listing), *host_is_local* (i.e. host lives in the same city as the listing), *host_multiple_listings* (i.e. the host has more than one listing) and *has_high_availability* (i.e. the availability of the listing that exceeds its city's short term rental period threshold: Amsterdam = 60 days, London = 90 days and Paris = 120 days).

The distance (i.e. *distance_cost*) from the city center is also calculated and identified in kilometers. Since the latitude and longitude coordinates are provided, the distance can be measured using the 'Haversine' formula, which gives the distance between two points on the surface of a sphere (Wang & Nicolau, 2017). The distance from each listing to the city is created in a way so that all the listings within the center have a distance of zero and the distance of listings falling outside this zone are equal to the total distance minus the radius of the circle that serves as a threshold (see Appendix F, Figures F.1 – F.3).

The additional control variables, i.e. Table 3, (i.e. *price_per_person*, *host_is_local*, *host_multiple_listings* and *has_high_availability*) are based on variables that were already provided in the datasets from Inside Airbnb. The host is a local (i.e. *host_is_local*) when the location of the host is the same as the listing location (e.g. both Amsterdam). Secondly, the *count of listings* is used to determine whether the host has more than one listing (i.e.). In addition, the estimated availability of the listing is used to determine whether or not a listing is *highly available* or not (i.e. *has_high_availability* =1= means that the city's maximum short-term rental period is exceeded) (see Appendix D, Table D.2).

Lastly, the natural logarithm is taken of the following variables: *price per person*, *security deposit*, *cleaning fee*, *extra people fee*, but also the *distance from the city center*. Since this research investigates the properties of the listing and therefore how it is perceived, the natural logarithm is taken of these numerical variables because people perceive these logarithmically (Dehaene, 2003). For example, if a person that is financially willing to pay 900 euros, 950 euros is not much more, while if you are only willing to pay 50 euros then the difference with 100 euros is perceived much more. Similarly, for *distance* the difference between being in the city center (i.e.

distance_cost = 0) and for example, 2 km outside of the city center (i.e. *distance_cost* = 2) is perceived as much further than the difference between for example 8 and 10 km.

3.3. Descriptive Statistics

Tables 4.1 – 4.3, below, show the descriptive statistics for each city. The number of observations differs per city, with Amsterdam having the smallest sample size of 3,909 observations as can be seen in Table 4.1. Every city has a similar spread in the dependent variable *occupancy_rate* with a standard deviation of around 0.23.

As mentioned, before to avoid certain "noise", the data per city is kept separately since each city has a different definition and requirements regarding short-term rental accommodations. For example, Amsterdam categorizes short-term rental periods to be shorter or equal to 60 days, London 90 days and Paris 120 days (see Appendix D, Table D.2). As can be seen in Tables 4.1 to 4.3, the mean occupancy rates in all three cities exceed the maximum short-term rental period (e.g. Amsterdam mean occupancy rate is 29%, which is equal to 105 days, exceeding the 60 days short-term rental period).

In addition, the presence of multicollinearity is also measured by using the variance inflation factor (VIF), which evaluates by how much the variance of the estimated regression coefficients was to increase in case the predictors are correlated (James et al., 2013). In the case of no multicollinearity, the variance inflation factor is equal to 1. A variance inflation factor between 5 and 10 is considered highly correlated, and above 10 is unacceptable (Craney and Surles, 2002). The correlation matrix showed overall low correlation (i.e. < 0.5), with the exception of the six Open badge scores amongst each other (i.e. < 0.6) (see Table 2), which are to be expected to positively correlate with each other). The variance inflation factor, as it calculates the correlation between a variable of interest and a group of variables, is a stronger indicator of multicollinearity in comparison to the correlation matrix, which solely takes into account bivariate relationships (James et al., 2013). Due to its low added value and very large table size, the correlation matrix is omitted from the appendices. The mean variance inflation factor lies closely to 1, which indicates that the predictors have relatively low and acceptable multicollinearity (see Appendix G, Table G.1).

Table 4.1. Sample characteristics for Amsterdam.

Variable	Obs.	Mean	Std. Dev.	Min.	Median	Max.
Occupancy rate (Y)	3,909	.291	.230	.003	.203	.7
*Cancellation policy: Moderate policy	3,909	.325	.468	--	--	--
* Cancellation policy: Strict policy	3,909	.489	.5	--	--	--
*Open Badge: Host is superhost	3,909	.292	.455	--	--	--
Open Badge: Score check-in	3,909	9.825	.525	2	10	10
Open Badge: Score communication	3,909	9.827	.571	2	10	10
Open Badge: Score accuracy	3,909	9.734	.629	2	10	10
Open Badge: Score value	3,909	9.219	.721	2	9	10
Open Badge: Score location	3,909	9.602	.634	2	10	10
Open Badge: Score cleanliness	3,909	9.537	.789	2	10	10
*Host gender	3,909	.522	.5	--	--	--
Price per person	3,909	59.642	28.380	5	55	250
Security deposit	3,909	168.579	353.954	0	0	4525
Cleaning fee	3,909	34.331	26.187	0	35	250
Extra people fee	3,909	13.406	25.845	0	0	272
Distance cost (km)	3,909	1.076	1.529	0	.546	9.220
*Instant bookable	3,909	.2957	.456	--	--	--
*Host is local	3,909	.847	.359	--	--	--
*Host multiple listings	3,909	.295	.456	--	--	--
*Has high availability	3,909	.669	.471	--	--	--
Accommodates	3,909	2.797	1.063	1	2	8

* For the dummy variables the following are not reported: minimum, median and maximum.

Table 4.2. Sample characteristics for Paris.

Variable	Obs.	Mean	Std. Dev.	Min.	Median	Max.
Occupancy rate (Y)	16,888	.366	.235	.003	.321	.7
* Cancellation policy: Moderate policy	16,888	.287	.452	--	--	--
* Cancellation policy: Strict policy	16,888	.459	.498	--	--	--
* Open Badge: Host is superhost	16,888	.245	.43	--	--	--
Open Badge: Score check-in	16,888	9.713	.667	2	10	10
Open Badge: Score communication	16,888	9.723	.677	2	10	10
Open Badge: Score accuracy	16,888	9.626	.752	2	10	10
Open Badge: Score value	16,888	9.245	.861	2	9	10
Open Badge: Score location	16,888	9.769	.569	2	10	10
Open Badge: Score cleanliness	16,888	9.283	.966	2	10	10
*Host gender	16,888	.537	.499	--	--	--
Price per person	16,888	40.375	21.546	2.25	35	258.75
Security deposit	16,888	340.379	531.579	0	200	4740
Cleaning fee	16,888	36	32.158	0	30	500
Extra people fee	16,888	7.746	15.869	0	0	269
Distance cost	16,888	.292	.563	0	0	6.701
*Instant bookable	16,888	.363	.481	--	--	--
*Host is local	16,888	.756	.429	--	--	--
*Host multiple listings	16,888	.335	.472	--	--	--
*Has high availability	16,888	.538	.499	--	--	--
Accommodates	16,888	3.063	1.451	1	2	16

* For the dummy variables the following are not reported: minimum, median and maximum.

Table 4.3. Sample characteristics for London.

Variable	Obs.	Mean	Std. Dev.	Min.	Median	Max.
Occupancy rate (Y)	25,873	.353	.237	.003	.308	0.7
* Cancellation policy: Moderate policy	25,873	.260	.439	--	--	--
* Cancellation policy: Strict policy	25,873	.509	.5	--	--	--
* Open Badge: Host is superhost	25,873	.265	.441	--	--	--
Open Badge: Score check-in	25,873	9.662	.851	2	10	10
Open Badge: Score communication	25,873	9.674	.862	2	10	10
Open Badge: Score accuracy	25,873	9.532	.958	2	10	10
Open Badge: Score value	25,873	9.283	1.001	2	9	10
Open Badge: Score location	25,873	9.596	.779	2	10	10
Open Badge: Score cleanliness	25,873	9.326	1.047	2	10	10
*Host gender	25,873	.547	.498	--	--	--
Price per person	25,873	34.402	21.317	1	30	299.5
Security deposit	25,873	128.663	267.889	0	0	4,009
Cleaning fee	25,873	32.202	33.512	0	25	610
Extra people fee	25,873	8.86	14.106	0	0	244
Distance cost	25,873	4.227	4.328	0	3.01	25.58
*Instant bookable	25,873	.430	.495	--	--	--
*Host is local	25,873	.655	.475	--	--	--
*Host multiple listings	25,873	.585	.493	--	--	--
*Has high availability	25,873	.643	.479	--	--	--
Accommodates	25,873	3.079	1.726	1	2	16

* For the dummy variables the following are not reported: minimum, median and maximum.

3.4. Model Specification and Analyses

The research investigates the relationship of strictness and quality with the performance of a listing by using an ordinary least squares regression model for the following cities: Amsterdam, Paris and London. Additionally, the moderator *host_gender* (i.e. 1= female host) is added to test whether or not the relationship between the performance (i.e. *occupancy_rate*) with the quality or the strictness of the listing is dependent on the gender of the host (Cohen et al., 2013) (see Figure 1).

3.4.1. Model Estimation

Before testing my hypotheses, to gain more insight into the dataset, the distribution and fit for each variable of the regression model are investigated (see Appendix H, Figures H.1.1–H.2.6). The regression model used to test the hypotheses is shown in the section below and is estimated using STATA. I use the results of the full model (i.e. Model 4) including the main and interaction effect of *host_gender* in order to test all my hypotheses.

3.4.2. Model Specification

The first hypothesis (i.e. H1.a and H1.b) tests whether or not the strictness of the cancellation policy signals of a listing has a destructive or beneficial effect on the performance of sharing rental accommodations. As aforementioned, contractual strictness can have two signaling effects: either signaling prestige/exclusivity and so attracting guests or signaling distrust and repelling guests (Ross, 1973; Spence, 1978; Zeithaml, 1988; Weiner, 2000; Srivastava and Lurie, 2001; Guo, 2009; Connelly et al., 2011; Oghazi et al., 2018). The dependent variable *occupancy_rate* is the proxy for performance and the *cancellation policy strictness* serves as a proxy for strictness. Additionally, to find out whether or not the gender of the host moderates the relationship between the performance and strictness (i.e. H1.c), the moderator is expressed in the form of interaction terms between the gender of the host (i.e. *host_gender*) and the cancellation policy strictness (i.e. *moderate_policy*host_gender* and *strict_policy*host_gender*).

Hypothesis 1:

H1.a. *A strict cancellation policy has a positive statistically significant effect on the occupancy rate.*

H1.b. *A strict cancellation policy has a negative statistically significant effect on the occupancy rate.*

H1.c. *The relationship between strict cancellation policy and the occupancy rate will be stronger among women.*

Furthermore, the second hypothesis tests if the quality signals provided on listing profiles have an effect on the performance. The dependent variable remains the *occupancy_rate* and the Open Badges are a proxy for quality indicators. In addition, the interaction terms between each badge and the gender of the host are included to investigate potential moderation (i.e. *host_gender*host_is_superhost*, *host_gender*score_communication_c9*, *host_gender*score_checkin_c9*, *host_gender*score_accuracy_c9*, *host_gender*score_value_c9*, *host_gender*score_location_c9* and *host_gender*score_cleanliness_c9*).

Hypothesis 2:
H2.a. *The Open Badges have a positive statistically significant effect on the occupancy rate.*
H2.b. *The relationship between Open Badges and the occupancy rate will be stronger among women.*

$$
\begin{aligned}
Y_{occupancy\_rate} \\
= \ & \alpha + \beta_1 X_{moderate\_policy} + \beta_2 X_{strict\_policy} + \beta_3 X_{host\_is\_superhost} \\
& + \beta_4 X_{score\_checkin\_c9} + \beta_5 X_{score\_communication\_c9} \\
& + \beta_6 X_{score\_accuracy\_c9} + \beta_7 X_{score\_value\_c9} + \beta_6 X_{score\_cleanliness\_c9} \\
& + \beta_8 X_{score\_location\_c9} + \beta_9 X_{host\_gender} + \beta_{10} X_{moderate\_policy} \\
& * X_{host\_gender} + \beta_{11} X_{strict\_policy} * X_{host\_gender} \\
& + \beta_{12} X_{host\_is\_superhost} * X_{host\_gender} + \beta_{13} X_{score\_checkin\_c9} \\
& * X_{host\_gender} + \beta_{14} X_{score\_communication\_c9} * X_{host\_gender} \\
& + \beta_{15} X_{score\_communication\_c9} * X_{host\_gender} + \beta_{16} X_{score\_accuracy\_c9} \\
& * X_{host\_gender} + \beta_{17} X_{score\_cleanliness\_c9} * X_{host\_gender} \\
& + \beta_{18} X_{score\_location\_c9} * X_{host\_gender} + \boldsymbol{\gamma' X} + \varepsilon_i
\end{aligned}
$$

, with $\boldsymbol{\gamma' X}$ captures all control variables (see Table 3) in a vector.

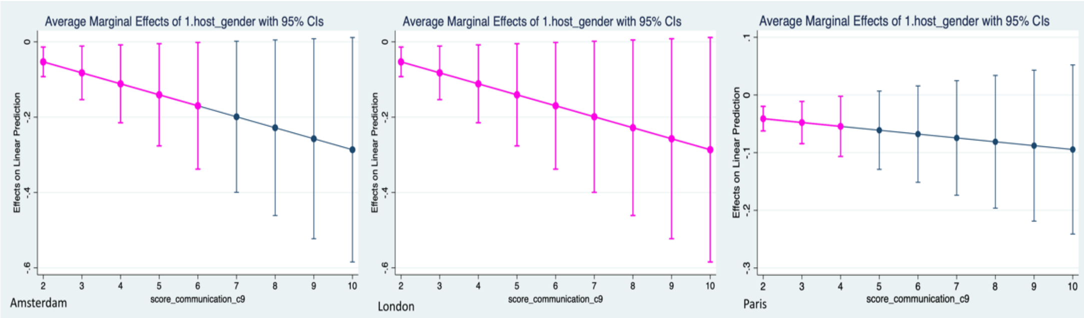
3.4.3. Moderating Effects: Mean-Centering and the Spotlight Analysis

To interpret the simple and interaction effects of the moderator *host_gender* with the scores that are Open Badges, the variables are shifted to a point of interest. In general, this is not needed for binary variables (i.e. *moderate_policy*, *strict_policy* and *host_is_superhost*) because 0 is already a value of interest (Spiller et al., 2013). These scores, as shown in Table 2, are discrete values ranging from 2 to 10. To enable the interpretation of the effect of gender of the host for the scores, the spotlight analysis is used (Spiller et al., 2013). The spotlight analysis provides the estimate of the simple effect of the variable *host_gender* at a selected level of the score Open Badges on the variable *occupancy_rate* (Spiller et al., 2013).

For the spotlight analysis, I choose as a relevant point the score level of 9. This point is chosen because from the analyses, the average in score Open Badge variables varies around 9 (see Table 4.1 – 4.3) and the frequency reveals that the discrete score of 9 is relatively more frequent (i.e. sometimes 10, but 9 is still chosen since the interest lies interpreting the effect of gender for an average score and 10 is the highest value) (see Appendix H, Figures H.3.1 – H.3.6). Therefore, the score value of 9 is chosen as a relevant point (i.e. new "0"). The centered score variable carries the "_c9" to indicate it is centered by 9 levels.

In the paper Spiller, Fitzsimons, Lynch and McClelland (2013), the spotlight analysis is visualized by showing the coefficient of the moderator for the selected value of an independent variable (i.e. Open Badges scores shown in Table 2). Here, following the idea of the floodlight method (i.e. performing the spotlight method for all points instead of focusing on a single point of interest) the margin plots are used to visualize the change in effect associated with the change of the variable *host_gender* (i.e. male versus female) at every level of a score Open Badge (i.e. a margin plot showing a significant negative value on linear prediction for a certain score level signifies that for that score level, being a female is related to having a lower occupancy rate than being a male).

Most of the marginal effects displayed in the margin plots are insignificant, there is however a common trend in which the marginal effect of variable *host_gender* is significant for the lowest score levels with the marginal value being negative (see Appendix E, Figures E.1 – E.5). This implies that a listing with low ratings and having a female host has on average a lower occupancy rate than a low rating listing having a male host. It is interesting to point out that the only score Open Badge for which the marginal effect is not only significant for low score levels is the variable *score_communication_c9*, most notably in London (see Figure 2). Communication requires a form of interaction between the host and the guest so it is expected that if the gender of the host were to have an effect that it would show more strongly in this score as opposed to dimensions unrelated to the *host_gender* (e.g. location).



Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure 2. Marginal effect of Host Gender on Occupancy Rate for all Scores of Communication, in Amsterdam, London and Paris.

4. Results

4.1. Hypotheses Testing

Table 5, shown below, provides an overview of the regression result of the full model (i.e. Model 4), the more extensive results per city can be found in the Tables I.1– I.3 in Appendix I.

Table 5. Overview Regression Results Model 4, with Dependent Variable Occupancy Rate, for Amsterdam, Paris and London.

Model 4 VARIABLES	Occupancy Rate Amsterdam	Occupancy_Rate Paris	Occupancy Rate London
Cancellation Policy (H1)			
moderate_policy	0.026**	0.023***	0.059***
strict_policy	0.018	0.054***	0.066***
Open Badges (H2)			
host_is_superhost	0.118***	0.121***	0.128***
score_checkin_c9	0.001	-0.007	-0.001
score_communication_c9	0.005	0.006	0.013***
score_accuracy_c9	-0.001	0.006	0.013***
score_value_c9	-0.013	-0.018***	-0.001
score_location_c9	0.021***	0.048***	0.016***
score_cleanliness_c9	0.01	0.001	-0.001
Moderators			
host_gender	-0.015	-0.028***	-0.009
host_gender*moderate_policy	0.013	0.003	-0.019**
host_gender*strict_policy	0.002	-0.014*	-0.026***
host_gender*host_is_superhost	0.016	0.005	0.015**
host_gender*score_checkin_c9	-0.001	0.006	0.0022
host_gender*score_communication_c9	-0.029*	-0.007	-0.007
host_gender*score_accuracy_c9	0.011	-0.001	-0.008
host_gender*score_value_c9	0.008	-0.003	0.002
host_gender*score_location_c9	0.004	0.0024	0.003
host_gender*score_cleanliness_c9	-0.002	0.000	-0.001
Control Variables			
log_price_per_person	-0.127***	-0.111***	-0.117***
log_security_deposit	4.48e-05	1.65e-05	-0.005***
log_cleaning_fee	-0.003	0.006***	-0.002**
log_extra_people_fee	-0.004**	-0.004***	-0.01***
log_distance_cost	-0.052***	-0.052***	-0.062***
instant_bookable	0.066***	0.097***	0.085***
host_is_local	0.002	-0.021***	-0.021***
host_multiple_listings	0.006	0.03***	0.003
has_high_availability	0.052***	0.040***	0.004
accommodates	-0.014***	-0.016***	-0.008***
neighborhood_n	-0.002***	0.001***	0.000
property_type_n	0.000	-0.000435	-0.001***
room_type_n	0.0867***	0.008***	-0.002
Constant	0.645***	0.664***	0.780***
Observations	3,909	16,888	25,873
R-squared	0.473	0.167	0.208
Adj. R-squared	0.468	0.165	0.207

NOTE: Variables 'log_variables' transformed to (natural) logarithms by means of the formula: log (variable+1), with +1 to avoid missing values.

Robust standard errors in parentheses

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

4.1.1. Cancellation Policy Strictness: Main Effects

Considering the first hypothesis (i.e.H1.a and H1.b), which looks at the *strict_policy*, I take into account the three types of cancellation policies (i.e. flexible, moderate and strict) with the *flexible_policy* as the reference category. Table 5, which summarizes model 4 for Amsterdam, Paris and London, shows that *moderate policy* has a positive statistical significant effect on the occupancy rate in all three cities (Amsterdam: $\beta = 0.026; p < 0.05$; Paris: $\beta = 0.023; p$< 0.05 and London: $\beta = 0.059; p < 0.05$). In addition, *strict_policy* has a positive significant effect on the occupancy rate in Paris (β =0.054; p <0.05) and London (β =0.066; ; p < 0.05) but not in Amsterdam (β =0.018; p >0.05). The Wald test demonstrates whether or not the variables *moderate_policy* and *strict_policy* significantly differ from each other (Gourieroux et al., 1982). The test results of each city are shown in Table 6 below. Since the significance level of the test is very close to 0 in each city (i.e. Amsterdam: 0.013, Paris and London: 0.000, p<0.05), I can reject the hypothesis of no difference between *moderate_policy* and *strict_policy*.

Table 6. Overview of Wald Test Results Comparing Variables *Moderate_Policy* and *Strict_Policy*, in Amsterdam, Paris and London.

Amsterdam		**Paris**		**London**	
F (2, 3877)	4.33	F (2, 16856)	58.13	F (2, 25841)	143.43
Prob > F	0.0133	Prob > F	0.000	Prob > F	0.000

These results show that a listing with a moderate policy leads to a higher occupancy rate on average. The findings suggest that moderate policy significantly increases the occupancy rate, suggesting the quality signaling effect of contractual strictness is present. With the exception of Amsterdam, strict policy is even a stronger quality signal than moderate policy. These findings suggest that strict cancellation policy has a significant positive effect on the occupancy rate, but that the effect size[1] is small, i.e. Table 7, (see Appendix H, Figures H.4.1 – H.4.9). The Cohen's d effect size estimates show that the cancellation policy strictness has a statistically significant,

[1] The effect size quantifies the size of the difference between two groups (Coe, 2002; Fritz et al., 2012). This is visualized in Appendix H Figures H.4.1 – H.4.9. The Cohen's d effect size estimate is used for the interpretation, for which a value of 0.8 represents a large effect size, 0.5 represents a medium effect size and 0.2 represents a small effect size (Fritz et al., 2012).

but small effect on the variable *occupancy_rate*[2]. Thus, while I do reject H1.b, I do not reject H1.a.

Table 7. Overview of the Effect Size Estimate Cohen's D for Amsterdam, Paris and London.

Variable	City	Cohen's D	95% C.I.
Cancellation policy: *Flexible_policy*	Amsterdam	0.035	-0.046; 0.115
	London	0.228	0.199; 0.257
	Paris	0.181	0.146, 0.215
Cancellation policy: *Moderate_policy*	Amsterdam	-0.018	-0.085; 0.049
	London	-0.066	-0.094; 0.038
	Paris	0.04	0.007; 0.074
Cancellation policy: *Strict_policy*	Amsterdam	-0.005	-0.068; 0.058
	London	-0.11	-0.134; -0.086
	Paris	-0.171	-.202; -0.141
Open Badge: *Host_is_superhost*	Amsterdam	-0.93	-1.002; -0.858
	London	-0.572	-0.599; -0.543
	Paris	-0.497	-0.532; -0.461
Moderator: *Host_gender*	Amsterdam	0.141	0.078; 0.203
	London	0.144	0.119; 0.168
	Paris	0.189	0.159; 0.219

[2] The effect size plots show that the occupancy rate distribution for each cancellation policy is similar except for two cases (see Appendix H, Figure H.4.1 – H.4.3). Firstly, places with a lower occupancy rate tend to opt for a flexible policy, which could be due to new listings with a low amount of reviews being more willing (or desperate) to improve the performance. Secondly, the spikes that can be seen for the highest occupancy rate show difference in distribution for the flexible and strict policies. This suggests that high performing places have a preference for a strict policy.

4.1.2. Cancellation Policy Strictness: Moderating Effect of Host Gender

I now test the last part of the hypothesis one (i.e. H1.c), i.e., whether or not being a female host amplifies the quality signaling effect of a strict cancellation policy (i.e. the effect of cancellation policy strictness on the occupancy rate). When looking at the results in Table 5, the significance of moderator *host_gender* differs per city. For Amsterdam, gender of the host has an insignificant effect (β = -0.015; p > 0.05). on the *occupancy_rate* and does not amplify the effect of a strict policy (β =0.002; p > 0.05). In Paris, the gender of the host has a negative statistically significant on the *occupancy_rate* (β = -0.028; p < 0.05), which means that being a female host has a negative effect on the performance of a listing. However, based on the interaction terms, being a female host in Paris does not amplify the effect of strictness on the performance (*host_gender*strict_policy*: β = -0.014; p > 0.05). Lastly, the findings in London show that the effect of the variable *host_gender* largely depends on the cancellation policy strictness (*moderate_policy*: β = -0.019 and *strict_policy*: β = -0.026; p < 0.05). This can be seen as the simple effect of host gender is insignificant, but the interactions of host gender with the various cancellation policies are significant.

The weak support, which is solely present in London, leads to the rejection of the third part of hypothesis one (i.e. H1.c). However, the inconsistency between cities suggests further investigation is needed.

4.1.3. Open Badges: Main Effects

The first part of the second hypothesis (H2.a) investigates the effect of the Open Badges indicating quality on the performance of a listing. The results in Table 5 show that being a superhost (*host_is_superhost*: Amsterdam: $\beta = 0.118$; Paris: $\beta = 0.121$ and London: $\beta = 0.128; p < 0.05$) and having a good location score (*score_location_c9*: Amsterdam: $\beta = 0.021; p < 0.05$; Paris: $\beta = 0.048; p < 0.05$ and London: $\beta = 0.016; p < 0.05$) has a positive statistically significant effect on the occupancy rate in all three cities. The effect size plots and Cohen' d in Table 7 comparing the superhosts versus non-superhosts, confirm these findings with large to medium effect sizes (see Appendix H, Figure H.4.4)

Table 5 shows that the remaining Open Badges are insignificant on an aggregated level. However, looking in detail the significance of the effect of a badge differs per city. Regarding Paris, the Open Badge score on value has a negative statistically significant effect (*score_value_c9*: β = -0.018; $p < 0.05$). A possible reason could be that Paris has a large city center, which is why some guests might have to compromise by accepting a lower price/quality ratio for a better location. On the other hand, the Open badges referring to score of communication with the host (i.e. *score_communication_c9*) and the score of accuracy of the listing description (i.e. *score_accuracy_c9*) are only (positive and statistically) significant in London (*score_communication_c9* β =0.013 and *score_accuracy_c9* β =0.013; $p < 0.05$). This suggests that guests value different listing qualities per city. Hence, due to the weak support (i.e. solely two out of seven are consistently significant), the first part of the second hypothesis (H2.a) is rejected.

4.1.4. Open Badges: Moderating Effect of Host Gender

The second part of hypothesis two (H2.b), investigates whether or not the effect of Open Badges on the occupancy rate is amplified by female hosts. The regression results from Table 5 show that none of the badges have a significant effect other than the host being a superhost and a female in London, which has a positive effect on the occupancy rate (*host_gender*host_is_superhost*: β = 0.015; $p < 0.05$). In the following graph, i.e. Figure 3, notice how the difference between the average *occupancy_rate* of female and male hosts is larger for non-superhosts (respectively the orange and red vertical lines) than for superhosts (the green and blue vertical lines). This shows that without the Open Badge superhost, the difference in performance is larger between male and female hosts in contrast to when they have proven their quality with the superhost badge.

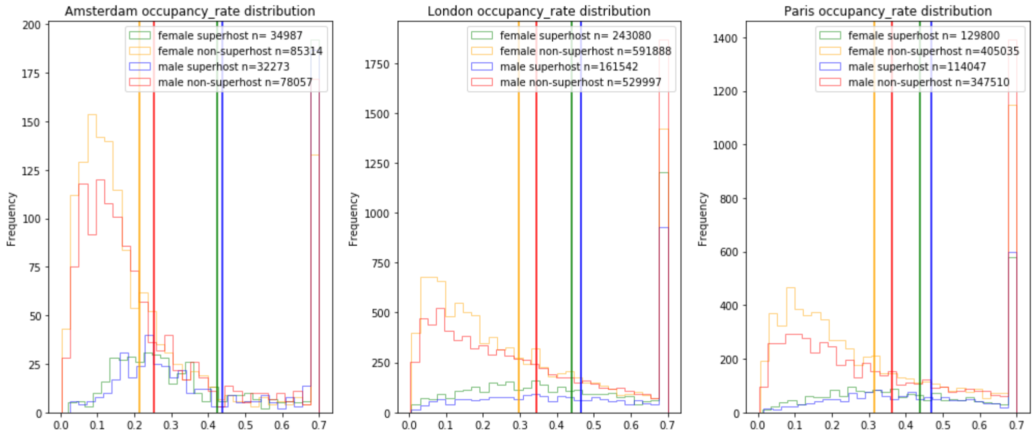


Figure 3. The Distribution and Means of *occupancy_rate* for Male and Female Superhosts and Non-Superhosts.

The margin plots give additional insight since these plots show the effect (coefficient and significant region) of host gender on occupancy rate for different score levels (Spiller et al., 2013). There is a strong trend in the values for which the marginal effects are significant: female hosts having low scores are penalized more on their performance as opposed to male hosts with low scores (see Appendix E, Figure E.1 – E.5). However, due to the weak support, the second part of hypothesis two (i.e. H2.b) is rejected.

Even though both hypotheses (i.e. H1 and H2) are rejected in general, with the exception of hypothesis 1.a, the significant effects from cancellation policy strictness and Open Badge types differ per city. The chosen cities are geographically close but are situated in different countries and have different cultural backgrounds. These cities are known to attract different types of tourists that may value different aspects of a listing (e.g. Paris is known as a romantic city). Amsterdam, a city that is known to be progressive and liberal, does not show signs of potential gender bias. However, the findings for Paris and London spark doubt regarding gender bias which deserves further investigation.

4.1.5. Control Variables

Several control variables are added to the regression model, which can be seen in Table 3. In general, the control variables are significant in most or all cities, here again, it is noticeable that the significant variables differ per city. One variable that stands out is the *instant_bookable* variable, which has a positive statically significant effect on the *occupancy rate* in all three cities (Amsterdam: $\beta = 0.0662; p < 0.05$; Paris $\beta = 0.0968.; p < 0.05$ and London $\beta = 0.08521; p < 0.05$). Additionally, the medium effect size (Amsterdam Cohen's d = 0.675; London Cohen's d = 0.378 and Paris Cohen's d = 0.473) confirms that instant bookable is a relatively important aspect of the performance of a listing (see Appendix H, Figure H.4.6). A listing that is instant bookable does not require a confirmation of the host before booking, which speeds up the booking process and allows the guest to complete the booking in one sitting.

5. Discussion

5.1. Conclusion with Implications for the Host and Airbnb

The first hypothesis tested the effect of contractual strictness proxied by the cancellation policy strictness. The first part (H1.a) stated that cancellation policy strictness helps the performance, which based on the results, I do not reject. Whereas the second part (H1.b) stated that cancellation policy strictness hurts the performance, I do reject. This means that listings that opt for a stricter cancellation policy have a better performance on average, with Amsterdam being the exception. The findings from the results suggest that there is indeed a prestige/exclusivity signaling effect of "strictness". In addition, the results showed solely weak proof of being a female host (negatively) amplifying these strictness signals.

The second hypothesis tested whether or not the Open Badges positively affect the performance of the listing. Unfortunately, also for this hypothesis solely weak support is found. There are only two Open Badges that consistently show positive significant effects on the performance, which is the Open Badge superhost and the score on the location. The superhost badge is an Open Badge that is specifically related to the quality of the host. The score of the location description is a more listing specific Open Badge as it entails the description of the location, but also the convenience of the location (see Table 2).

Even though both hypotheses are rejected, the significant results in London spark doubt as these findings imply that when a female host is strict, her listing's performance is more hurt than if it were to have a male host. On the other hand, the results also show that when a female host is a superhost, her listing's performance improves more than if the superhost were to be a male, due to the larger difference in performance between the genders for non-superhosts. This could indicate that when a guest has less quality signaling attributes to rely on, the remaining discerning attributes such as the gender of the host have a larger weight. Therefore, even though the hypotheses were rejected in general, the inconsistency in significant variables per city shows that further research is needed.

The results confirm that strictness positively affects the perceived quality, which is prestige/exclusivity (Hjorth-Anderson, 1984; Monroe and Krishnan, 1985; Zeithaml, 1988). However, the Open Badges in general, which are supposed to signal the objective quality as it is measurable and verifiable, do not necessarily improve the performance except for the Open Badge superhost.

These findings bring up implications for the hosts and Airbnb. The implication for the host is that being strict is to the hosts' advantage (i.e. in case the guest cancels) and helps performance according to the results. Additionally, most listings have high review scores so these are less discerning and should not be relied on fully, however, low scores are rare therefore a more discerning signal and related to low occupancy. Therefore, if the quality indicators are not discerning enough, guests need to rely on other characteristics of the listing and host profile, which may cause a gateway to discrimination from guests towards hosts.

Listing attributes valued by guests may depend on geography and culture, but also the aim of the vacation. This is an implication for Airbnb, being the largest sharing rental accommodation platform, to provide guests a way to distinguish the listings and help the hosts compete more fairly. Furthermore, previous studies have already pointed out the overestimation of the score Open Badges (also referred to as ratings) (Zervas et al., 2015; Salganik et al. 2006). In this research, besides the superhost Open Badge, the scores do not appear to be clear indicators of good performance.

5.2. Limitations and Future Research

5.2.1. Limitations

The findings from the first hypothesis suggest a prestige/exclusivity signaling effect from strictness, however, when looking deeper, there is a possibility that this is a case of reverse causality. Reverse causality is the case when the direction of the relationship between two variables is opposite to what is expected (Chong and Calderon, 2000). For example, that being a better performing listing causes it to be stricter as opposed to that being contractually strict improves the listing's performance.

The effect size plots (see Appendix H, Figures H.4.1 – H.4.9) show that there is no big difference in the distribution of performance between the strict cancellation policies versus the non-strict cancellation policies, except for the high performing listings (i.e. maximum occupancy rate of 70%). This could suggest that when listings are among the better performing ones, hosts then opt for stricter cancellation policy as they have established a certain reputation through earning Open Badges and good reviews.

In general, reverse causality is a big limitation to this research since the research is based on data from one point in time (i.e. 9th of May 2020). This means that the tests could not verify influences from the listings' performance in the past. This can be done by collecting several datasets of the same cities matching the same listings over various points in time and comparing the changes in strictness and performance over time. For example, by using a difference-in-difference model, which is a version of a fixed-effects estimation by using aggregate data and could look at the listings' change in cancellation policy strictness (Angrist and Pischke, 2008).

The second limitation is the presence of omitted variable bias (Angrist and Pischke, 2008). This means that relevant variables needed were missing in the regression models. As mentioned before, besides missing data from the time periods before, the actual availability was missing, which is needed to estimate the occupancy rate. This research had no option but to rely on the estimated availability of this year 2020. Since it is estimated, it means that hosts can change the availability of the listing at any moment. Missing the availability also causes that the results do not take into account a certain unfairness in the comparison of performance. Listings that respect the short-term rental period, e.g. in Amsterdam 60 days, and are fully booked, should be considered high performers. However, since the availability is missing, the used denominator is 365 days, which makes them seem as low performers as opposed to listings that exceed the short-term rental period, i.e. > 60 days.

Furthermore, the dependent variable occupancy rate was estimated based on the Inside Airbnb occupancy model (see Appendix D, Table D.1). This could have been estimated wrongly for hosts that have many listings. Hosts that have more listings could cause them to be relatively less personal as opposed to hosts that host occasionally, and so may generate fewer reviews per booking on average, which will affect the estimation of the occupancy rate.

The findings provide quite some support for contractual strictness having a prestige/exclusivity signaling effect and not a repelling effect. However, both signaling effects could be at play with the repelling effect only kicking in later for higher levels of strictness. The original dataset contained two more levels of strictness (i.e. super strict), which were also the strictest, but due to small sample size were merged with strict policy (see Table 1). Having more data for these highest levels of strictness, one could keep these "super strict" cancellation policies separate to possibly gain more insight into the repelling effect.

Lastly, of all Open Badges, superhost was the only one to stand out as an important driver of performance. However, the Open Badge scores, which are presented as continuous variables on a scale of five were changed to discrete values on a scale from two to ten. This caused the loss of better distinction among the listings.

5.2.2. Future Research

For further research, it would be interesting to compare the research using the actual occupancy rate and testing the data over time using models such as a difference-in-difference model to confirm findings. In addition, other variables could be added, for example the host types which can be gathered by conducting interviews with hosts (e.g. professional and non-professional hosts) and/or use text analysis (e.g. performing sentiment analysis on the guest reviews and host answers).

Besides the supply side of the sharing rental accommodation market, the research is missing the demand side. This could be an additional possibility for further expansion of this research. This could be done through for example conducting a survey amongst (potential) guests. The survey could help with answering the questions in a better way and with more quality assurance. Besides the gender, race or skin color could be included in order to better understand or maybe unveil more discrimination. For example, the research could find out what order guests decide on how to influence or nudge guests into making less discriminatory choices.

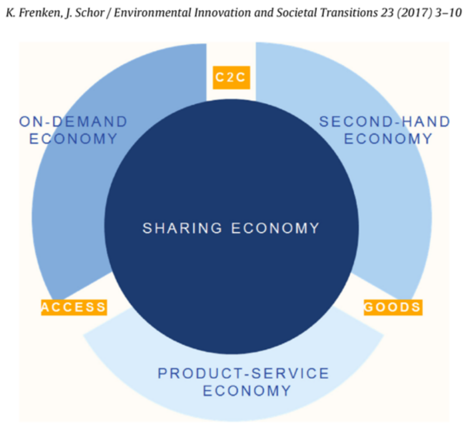
Current events regarding the Trump administration and the largest civil rights movement led by Black Lives Matter (i.e. protests were worldwide) have brought up greater discussion about discrimination and racism (Buchanan et al., 2020).

Companies, such as Goya Foods, that have publicly supported and praised president Trump are facing calls for boycotts (McClay, 2020). Racism on Airbnb has already led to the platform "Noirbnb", which provides "a better, safer experience for travelers of color" (https://noirbnb.com/about). Further research could help in understanding what improves performance and what may cause the gateway to discrimination. For example, Airbnb in order to improve its status as "the" platform, could actively try to prevent discrimination. The better the Open Badges the fewer guests need to rely on basic characteristics such as gender or skin color.

As a final note, while writing this paper, the COVID-19 pandemic started, which some say will change our lifestyle. It currently has changed how educational institutions and companies need to operate, for example restaurants needing to quickly switch to delivery (Power, 2020). The results show that the Open Badges in form of scores on for example cleanliness to be insignificant for the performance of a listing. The spread of the pandemic might make people think more about this, potentially altering the effect of the cleanliness badge.

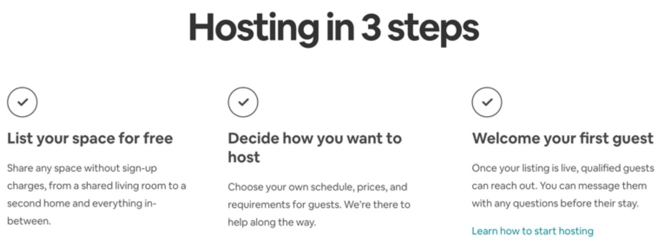
Further research could investigate if specificity and the context of Open Badges play an important role in their relation to a listing`s performance and in their relation to discrimination.

Appendix A
The Sharing Economy: Overview



Source Frenken et al., 2015

Figure A.1. The Sharing Economy and Related Forms of Platform Economy.

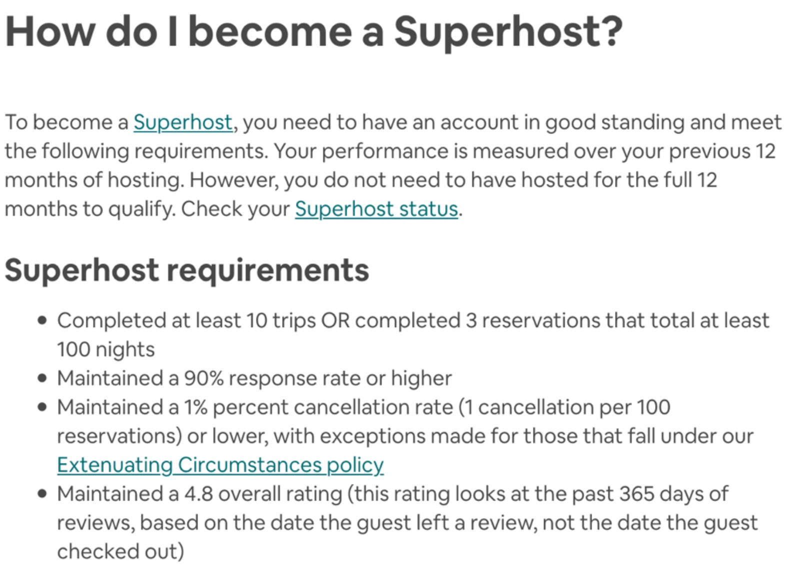


Source Airbnb.com

Figure A.2. Three steps to become a host on Airbnb.com.

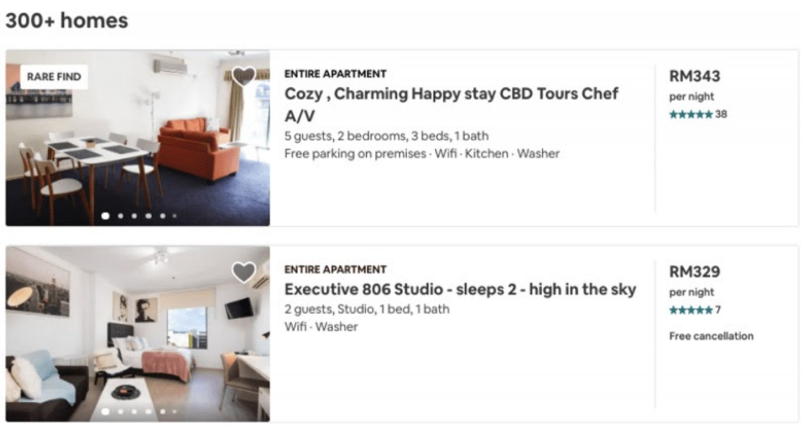
Appendix B

Open Badges: Description and Requirements



Source Airbnb.com

Figure B.1. The requirements of the Open Badge "Superhost".



Source Airbnb.com

Figure B.2. Example of Open Badge "Rare find".

Appendix B continued

How can my listing qualify for the Airbnb Plus program?

The Airbnb team reviews current listings for exceptional quality and design and invites hosts with eligible listings to join the program.

Requirements

To be eligible for an invitation to Airbnb Plus, listings must be located in areas we support and hosts must demonstrate Superhost-level hospitality and have:

- Maintained an average rating of 4.8 over the past year
- No canceled reservations over the past year (unless there were extenuating circumstances)

Airbnb Plus hosts are also held to additional exceptional hospitality standards by demonstrating that they:

- Genuinely care by welcoming guests warmly and delighting them with details that let them know they've thought of everything
- Provide outstanding service by being kind, responsive, and committed to making things right

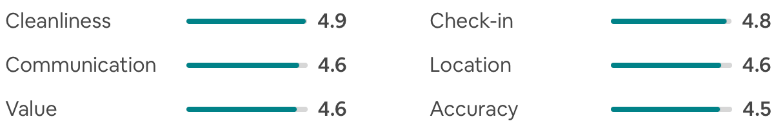
Additionally, Airbnb Plus spaces must be:

- Well-designed in a way that shows the host's unique style and personality
- Fully equipped with amenities like wifi, coffee, and quality linens
- Spotlessly maintained, extra clean, and clutter-free

Read the Airbnb Plus Terms and Conditions to find out more.

Source Airbnb.com

Figure B.3. The requirements to obtain Open Badge "Plus".



Source Airbnb.com

Figure B.4. Illustration of Open Badges on the listing profile.

Appendix B continued

Guests can give ratings on:

- **Overall experience.** Overall, how was the stay?
- **Cleanliness.** Did guests feel that the space was clean and tidy?
- **Accuracy.** How accurately did the listing page represent the space? For example, guests should be able to find up-to-date info and photos in the listing description.
- **Value.** Did the guest feel that the listing provided good value for the price?
- **Communication.** How well did you communicate before and during the stay? Guests often care that their host responds quickly, reliably, and frequently to their messages and questions.
- **Check-in.** How smoothly did check-in go?
- **Location.** How did guests feel about the neighborhood? This may mean that there's an accurate description for proximity and access to transportation, shopping centers, city center, etc., and a description that includes special considerations, like noise, and family safety.
- **Amenities.** How did guests feel about the amenities that were available during their stay? Guests often care that all the amenities listed are available, working, and in good condition.

In each category, hosts are able to see how often they get 5 stars, how guests rated nearby hosts, and, in some cases, tips to help improve the listing.

The number of stars displayed at the top of a listing page is an aggregate of the primary scores guests have given for that listing. At the bottom of a listing page there's an aggregate for each category rating. A host needs to receive star ratings from at least 3 guests before their aggregate score appears.

Source Airbnb.com

Figure B.5. Explanation of Airbnb.com rating system.

Appendix C

Literature Table

Table C.1. Literature Table

Author(s), Year	Empirical/Theoretical approach	Key Findings
Brucks et al., 2000	▪ **Qualitative study** Two focus groups and interviews ($N_{women} = 10; N_{Men} = 10;\ N_{DurableGoods} = 10$) Telephone interviews ($N_{MarketingResearchManager} = 3$) Previous study by Brucks (1985) ($N_{Attributes} = 100; N_{Consumers} = 36$) ▪ **Process-tracing laboratory experiment** ($N_{women} = 50; N_{Men} = 50$)	▪ The properties of quality dimensions: search, experience and credence ▪ Consumers use price and brand name to judge prestige
Dellaert, 2019	**Theoretical approach**: ▪ Two-layered conceptual framework of consumer co-production taking into account co-production activities. ▪ Combination of household production theory with institutional design theory and consumer behavior.	▪ Consumers in a peer-to-peer market are co-producers ▪ Supportive marketing at the consumer co-production level are divided in two categories: 1. Offering collective co-production services (e.g. matching algorithms, collective insurances) 2. Establishing clear rules of engagement w.r.t. within-network market behavior (e.g. consumer co-production quality measures, reward mechanism)
Fornell et al., 1996	**Theoretical approach**: ▪ American Customer Satisfaction Index (ASCI) ($N_{Firms} > 200, N_{Industries} > 40$ in 7 major consumer sectors (e.g. Retail, Manufacturing/Durables, etc.)	▪ Customization is more important than reliability in determining customer satisfaction. ▪ Customer expectations play a great role in sectors in which variance in production and consumption is relatively low. ▪ Customer satisfaction is more quality-driven than value-or price-driven.
Guo, 2009	**Theoretical approach:** ▪ Expansion of previous analytical research with emphasis on competition (Xie and Gerstner, 2007). ▪ Questions (1) How would a partial refund policy influence the firms' strategic interaction? (2) How sustainable are partial refund policies in a competitive market?	▪ Advance selling (i.e. reserving) is common in service industries that appear highly competitive. ▪ Partial refunds can endogenously change the nature of strategic interaction between service providers from local monopolies into a competition regime. ▪ Partial refunds are efficiency improving and competition intensifying.
Huang & Dev, 2019	**Secondary data**: ▪ Two longitudinal datasets merged together (BAV consulting, division of Young&Rubicam)($N_{ServiceBrand\_year} = 5634;\ N_{UniquePublicServiceFirm} = 502$).	▪ Three drivers service brand growth: quality, personalization and relationships. ▪ Relationship-based service personalization is advised (as opposed to quality-based service) to maintain quality at customer expectations.

Jiang & Rosenbloom, 2005	**Secondary data:** ▪ BizRate.com survey ($N_{Respondents} >$ 250,000; $N_{e-tailers} = 416$), June 2002	▪ After-delivery satisfaction has a greater influence on both overall customer satisfaction and intention to return than at check-out satisfaction. ▪ Price perception, measured on comparative basis, has a direct and positive effect on overall satisfaction and intention to return.
Li & Srivasan, 2019	**Secondary data**: ▪ Listing-level Airbnb data ($N_{Properties} =$ 342,873), August 2014 - October 2015 ▪ Hotel data (collected by Smith Travel Research; ($N_{Hotels} = 4{,}943$), January 2008 – October 2015	▪ The entry of flexible-capacity sharing platforms (e.g. Airbnb, Uber) affects the competitive landscape in traditional industries with fixed capacity and demand. ▪ Airbnb midly cannibalizes hotel sales and expand market for hospitality industry.
Liu & Arnett 2000	▪ **Survey** ($N_{Fortune1000Companies} =$ 762; $N_{Webmasters} = 689$)	Four factors critical to Web site success in e-commerce: ▪ Information and service quality ▪ System use ▪ Playfulness ▪ System design and quality
Oghazi et al., 2018	**Secondary data:** ▪ ($N_{OnlineConsumers} = 730$),	▪ Return policy as a market signaling mechanism is a costly investment to not only support current transactions, but also to signal commitment towards customer service.
Pansari & Kumar, 2017	**Theoretical approach**: ▪ Based on marketing literature and popular press articles. ▪ Develop a framework for customer engagement.	▪ When the relationship with consumer entails satisfaction and emotional connectedness, partners become more concerned about each other (engagement-focused).
Simonsohn, 2010	**Second data:** ▪ eBay ($N_{BestSellingSingleDVD\_Auctions} =$ 11,796; $N_{MultipleDVDBundles\_Auctions} =$ 3,177),	Consistent with competition neglect, it is found that: ▪ A disproportionate share of auctions end during peak bidding hours. ▪ Such hours exhibit lower selling rates and prices. ▪ Peak listing is more prevalent among sellers likely to having chosen ending time strategically, suggesting disproportionate entry is a mistake driven by bounded rationality rather than mindlessness.
Srivastava & Lurie, 2001	**Field study**: ▪ Study 1 $N_{UniversityStaffMembers} = 146$ ▪ Study 2 $N_{Undergraduates} = 69$	▪ In retail, the presence of a refund increases likelihood of discontinuing search.
Voss et al., 1998	**Experiment**: ▪ Simulation of hotel service exchange, collect pre-and post-purchase measures ($N_{Subjects} = 200$)	▪ Post-purchase price perception has a major impact on satisfaction almost as high as that of performance perceptions. ▪ Post-purchase performance and price perceptions significantly influence satisfaction.
Xie & Gerstner, 2007	**Theoretical approach**	- Offering refunds for service cancellation can be profitable: ▪ Without charging a higher price compared with a no refund. ▪ Even when advance buyers would want to cancel the service. ▪ Service providers should decrease the "hassle cost" of cancellations. - Profit advantage of advance selling, i.e. captures consumer-added surplus created when customers find

		new alternatives. The customers are therefore more willing to pay a fee to terminate the pre-purchased contract (i.e. partial refunds). -Offering refunds for cancellation reduces the need to reserve capacity for high-paying customers and improves capacity utilization.
Xie et al., 2014	**Field study** ($N_{Hotels} > 843$)	Significantly associated with hotel performance: ▪ Overall rating, attribute ratings of purchase value. ▪ Location and cleanliness. ▪ Variation and number of consumer reviews. ▪ Number of management responses.
Zervas et al., 2017	▪ **Field study**: Texas (Airbnb.com) 2008-2013 ($N_{Listings} > 22{,}000$). ▪ **Secondary data**: quarterly hotel revenue tax data 2003-2013 ($N_{Hotels} = 4{,}000$).	▪ Consumer supplied accommodations lower hotel room prices. Main reason consumer supply more flexible.
Zeithaml, 2000	**Theoretical approach**: ▪ Conceptual framework linking service quality (split up into 6 categories) in relation with profits.	▪ The customer perceptions of the service quality affect the purchase intentions. ▪ The key drivers of: service quality, customer retention and profits.

Appendix D

The Performance of Sharing Rental Accommodations

Table D.1. The Occupancy Model: "San Francisco Model".

Occupancy Model	Calculation	Reason
Review rate of 50%	(Reviews per month x 12 (months)) / 0.5	Convert reviews to **estimated # bookings**. Why 50%? - Airbnb stated that 72% of the guests leave a review (1). However, this is unverifiable (i.e. need an API key for this/confidential). - The B&L Analyst Office use 72%, but also introduce higher impact model with a 30.5% review rate (based on public data of reviews (2)(3). - InsideAirbnb.com found a review rate of 30.5% more fact based but not conservative enough (i.e. B&L Analyst Office did not take into account listings that were taken off or missing reviews). - 50% is chosen as it almost exactly lies between 72 and 30.5%
Average length of stay	London – 5.2 nights Amsterdam – 3.9 nights Paris – 4.6 nights	- The average stay for the chosen cities, which is publicly reported (Airbnb Economic Impact, 2015). - However, this is adapted for listings where the number of minimum nights exceeds this average.

Estimated # bookings x **Average length of stay** = occupancy days/365 = **Occupancy rate**

Cap occupancy rate at 70%	N/A	Why? - 70% is a realistic number for highly occupied hotels (Priyadarsini et al., 2009). - Controls for situations such as hosts changing the stated minimum nights during high season (i.e. review data is not established yet). - To keep the occupancy model conservative.

(1) Executive Summary of Amendments Relating to Short-Term Rentals https://commissions.sfplanning.org/cpcpackets/2014-001033PCA.pdf
(2) Budget and Legislative Analyst – Analysis of the impact of short-term rentals on housing https://sfbos.org/sites/default/files/FileCenter/Documents/52601-BLA.ShortTermRentals.051315.pdf
(3) Attorney General and City of New York announce Joint Enforcement Initiative Against Illegal Hotels. https://ag.ny.gov/pdfs/AIRBNB%20REPORT.pdf

Appendix D continued

Table D.2. Short-term Rental Requirements for Amsterdam, London and Paris.

City	Short-term rental threshold (maximum availability)	Requirements
Amsterdam	60 days	- Only the registered occupants are allowed to rent out property, however registered tenants are not allowed. - The maximum time an Airbnb host in Amsterdam can rent out his/her place is 2 months (60 days) to a maximum of 4 people at a time. - If the threshold is exceeded, the owner needs to register as entrepreneur for tax purposes.
London	90 days	- Property owner needs a "planning permission" if 90 days are exceeded otherwise fine of £20,000. (1) - Planning permission is needed when owner: (2) • builds something new • makes a major change to the building (e.g. extension to building) • or changes the use of the building.
Paris	120 days	- Property owners are allowed to rent their primary residence (i.e. only if owner lives there) for a maximum time of four months. - It is illegal to rent out property for less than a year at a time if the owner does not live there permanently. However, if registered and in the possession of a license as a commercial property with the city it is legal. - If caught renting out unlicensed secondary apartments or renting out their primary residences for more than four months are obliged to pay a fine of € 25,000. (3)

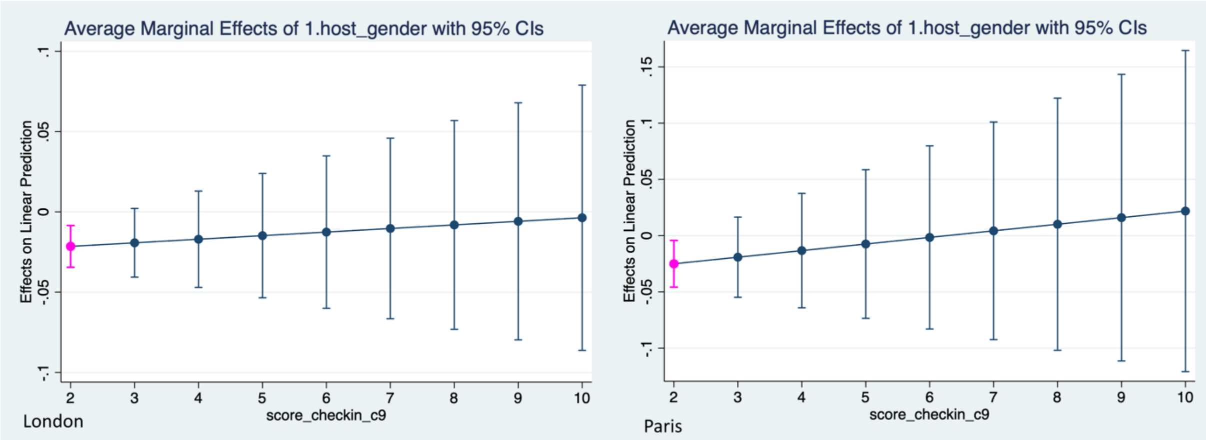
(1)https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/belastingdienst/zakelijk/internationaal/btw_voor_buitenlandse_ondernemers/onroerende_zaken/verhuren/verhuur_vakantiewoning

(2) Section 25 of the Greater London Council (General Powers) Act 1973 (as amended by Section 4 of the Greater London Council (General Powers) Act 1983). http://www.legislation.gov.uk/ukpga/2015/20/notes/division/5/46

(3) https://www.gov.uk/planning-permission-england-wales

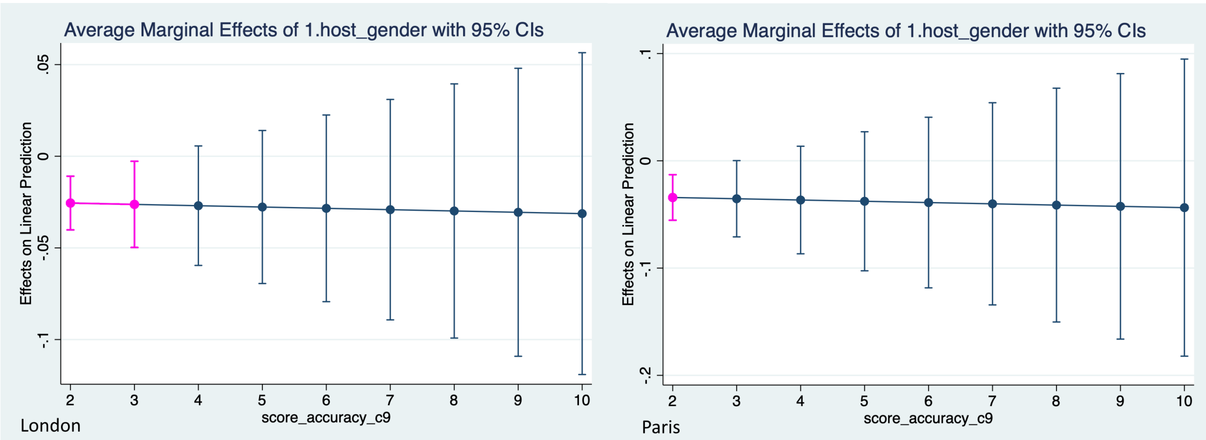
(4) http://www.splm-france.fr/en/proposals-for-housing-bill-alur/

Appendix E
The Marginal Effects of Host Gender on Occupancy Rate for the Score Open Badges



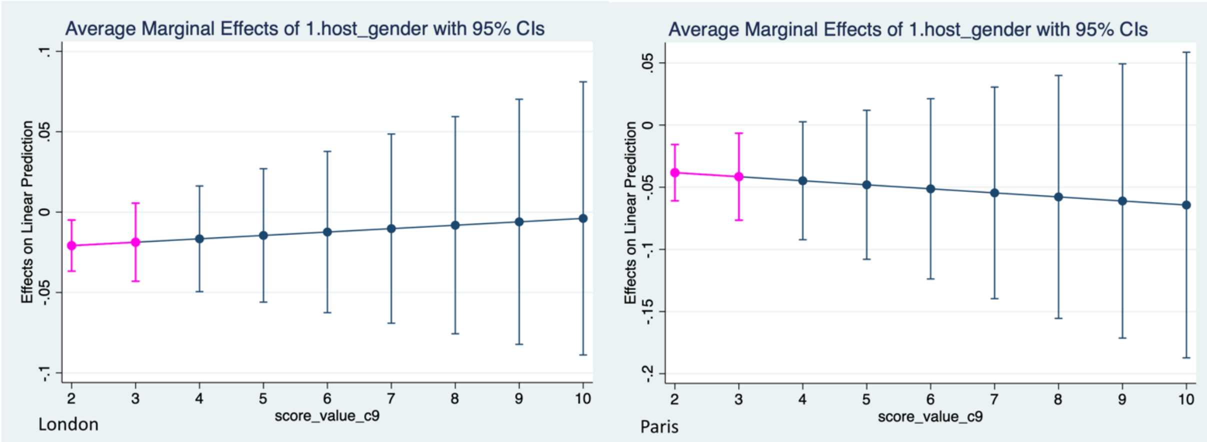
Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure E.1. Marginal effect of Host Gender on Occupancy Rate for all Scores of Check-in, in London and Paris.



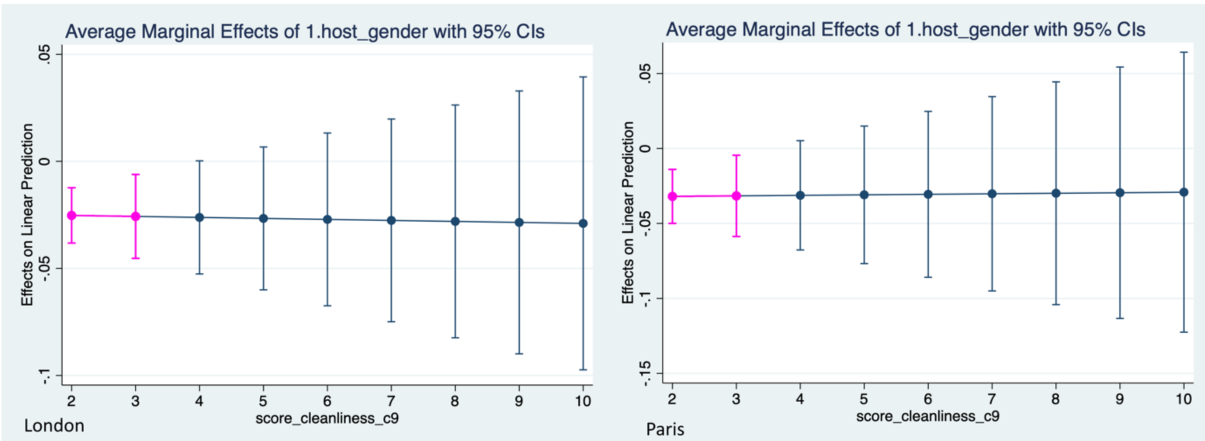
Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure E.2. Marginal effect of Host Gender on Occupancy Rate for all Scores of Accuracy, in London and Paris.



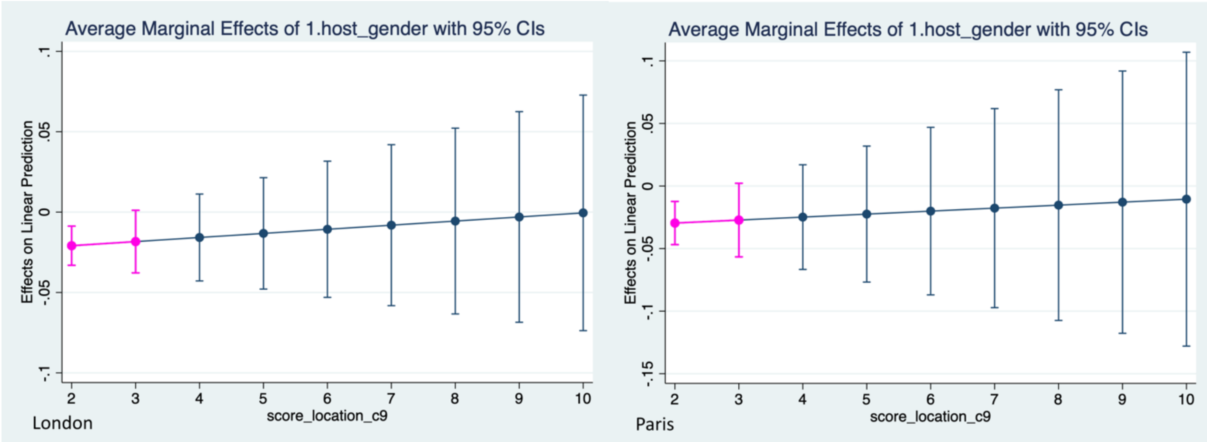
Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure E.3. Marginal effect of Host Gender on Occupancy Rate for all Scores of Value, in London and Paris.



Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure E.4. Marginal effect of Host Gender on Occupancy Rate for all Scores of Cleanliness, in London and Paris.

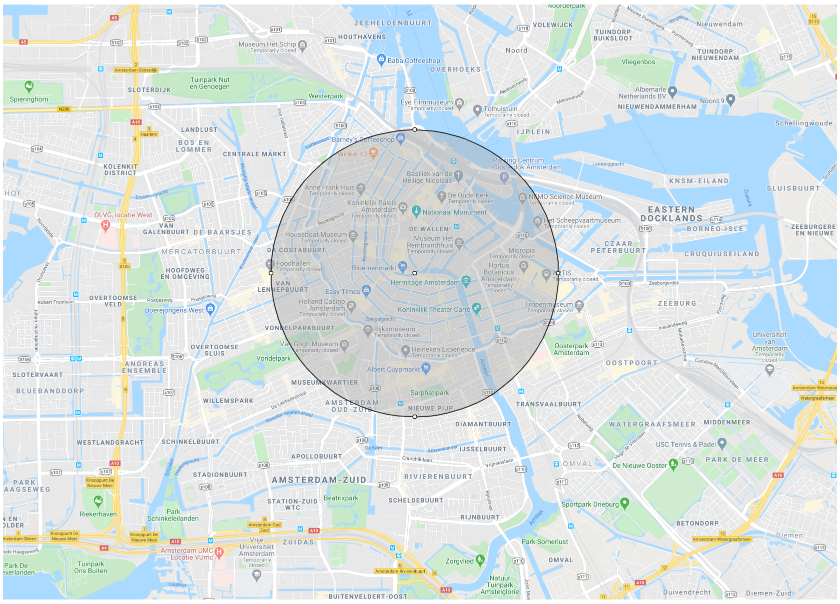


Note: area colored in purple is the region of the Open Badge score on the x-axis, where the variable *host_gender* is significant ($p < 0.05$).

Figure E.5. Marginal effect of Host Gender on Occupancy Rate for all Scores of Location, in London and Paris.

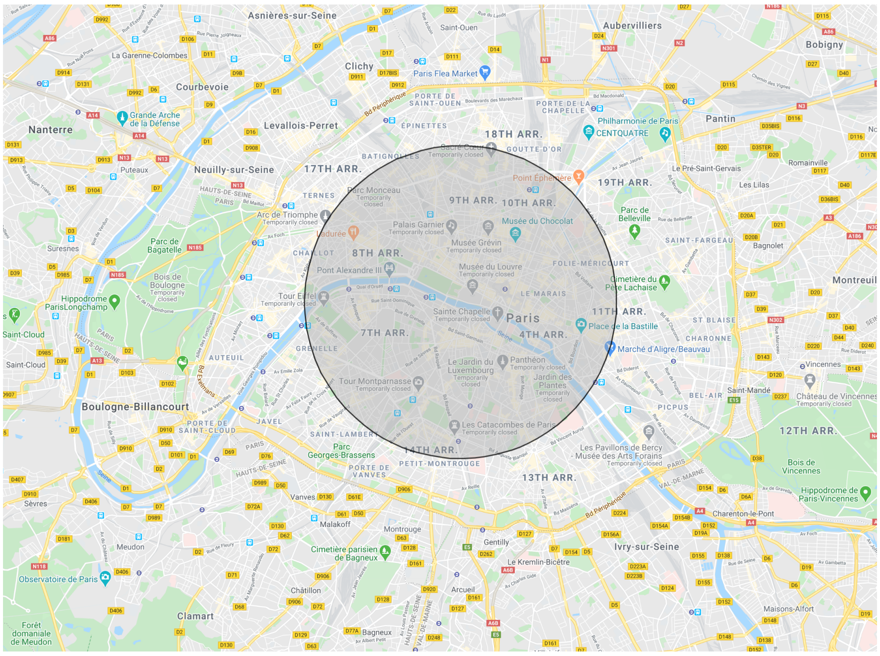
Appendix F

City Centers in Amsterdam, Paris and London



Radius of Center = 1.71 km; Coordinates – position: 52.367097,4.893426

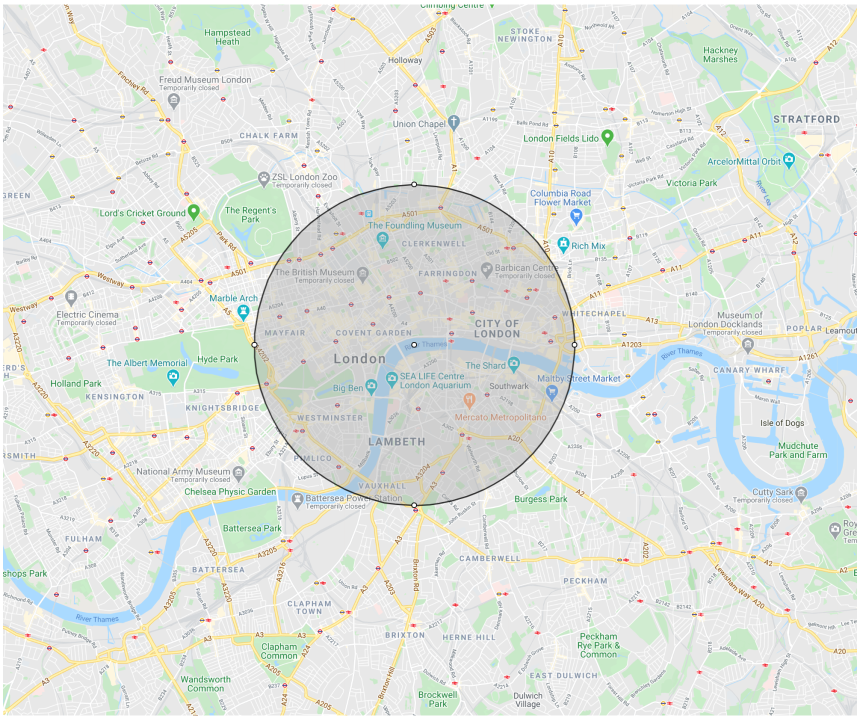
Figure F.1. Illustration of Determined Central Zone for Amsterdam.



Radius of Center = 3.32 km; Coordinates – position: 48.748294, 2.605041

Figure F.2. Illustration of Determined Central Zone for Paris.

Appendix F continued



Radius of Center = 2.97 km; Coordinates – position: 51.509565, -0.113057

Figure F.3. Illustration of Determined Central Zone for London.

Appendix G

Multicollinearity Measure

Table G.1. Variance Inflation Factor for Amsterdam, Paris and London.

Amsterdam			Paris			London		
Variable	**VIF**	**1/VIF**	**Variable**	**VIF**	**1/VIF**	**Variable**	**VIF**	**1/VIF**
Score_communication	2.47	0.404	Score_accuracy	2.27	0.441	Score_accuracy	3.16	0.317
Score_accuracy	2.36	0.424	Score_value	2.20	0.454	Score_value	3.02	0.331
Score_checkin	2.10	0.476	Score_communication	2.15	0.465	Score_communication	2.58	0.388
Strict_policy	2.03	0.494	Score_checkin	1.99	0.502	Score_checkin	2.39	0.419
Score_cleanliness	1.97	0.507	Score_cleanliness	1.86	0.539	Room_type_n	2.30	0.435
Moderate_policy	1.95	0.512	Strict_policy	1.70	0.589	Score_cleanliness	2.16	0.463
Score_value	1.88	0.533	Moderate_policy	1.58	0.631	Accomodates	1.97	0.507
Room_type_n	1.75	0.572	Log_cleaning_fee	1.37	0.730	Strict_policy	1.75	0.570
Score_location	1.71	0.584	Score_location	1.35	0.738	Score_location	1.65	0.605
Log_distance_cost	1.66	0.602	Accommodates	1.30	0.772	Moderate_policy	1.64	0.610
Log_price_per_person	1.53	0.654	Log_security_deposit	1.28	0.778	Log_cleaning_fee	1.57	0.638
Accommodates	1.40	0.716	Log_price_per_person	1.26	0.795	Log_price_per_person	1.52	0.658
Log_cleaning_fee	1.22	0.816	Room_type_n	1.22	0.819	Log_distance_cost	1.46	0.686
Log_security_deposit	1.19	0.843	Host_multiple_listings	1.16	0.859	Log_security_deposit	1.36	0.733
Host_is_superhost	1.18	0.848	Host_is_superhost	1.12	0.889	Property_type_n	1.22	0.817
Instant_bookable	1.17	0.857	Has_high_availability	1.12	0.894	Host_is_superhost	1.13	0.888
Host_multiple_listings	1.16	0.866	Log_extra_people_fee	1.12	0.895	Log_extra_people_fee	1.12	0.890
Log_extra_people_fee	1.13	0.884	Instant_bookable	1.12	0.911	Host_multiple_listings	1.12	0.893
Property_type_n	1.13	0.887	Log_distance_cost	1.09	0.917	Neighborhood_n	1.08	0.924
Neighborhood_n	1.11	0.899	Neighborhood_n	1.07	0.931	Instant_bookable	1.08	0.926
Has_high_availability	1.08	0.926	Host_is_local	1.04	0.959	Host_gender	1.04	0.958
Host_is_local	1.05	0.953	Property_type_n	1.04	0.965	Host_is_local	1.04	0.959
Host_gender	1.04	0.965	Host_gender	1.03	0.969	Has_high_availability	1.04	0.960
Mean VIF	**1.53**		**Mean VIF**	**1.41**		**Mean VIF**	**1.67**	

Appendix H
Fit and Distribution of Variables

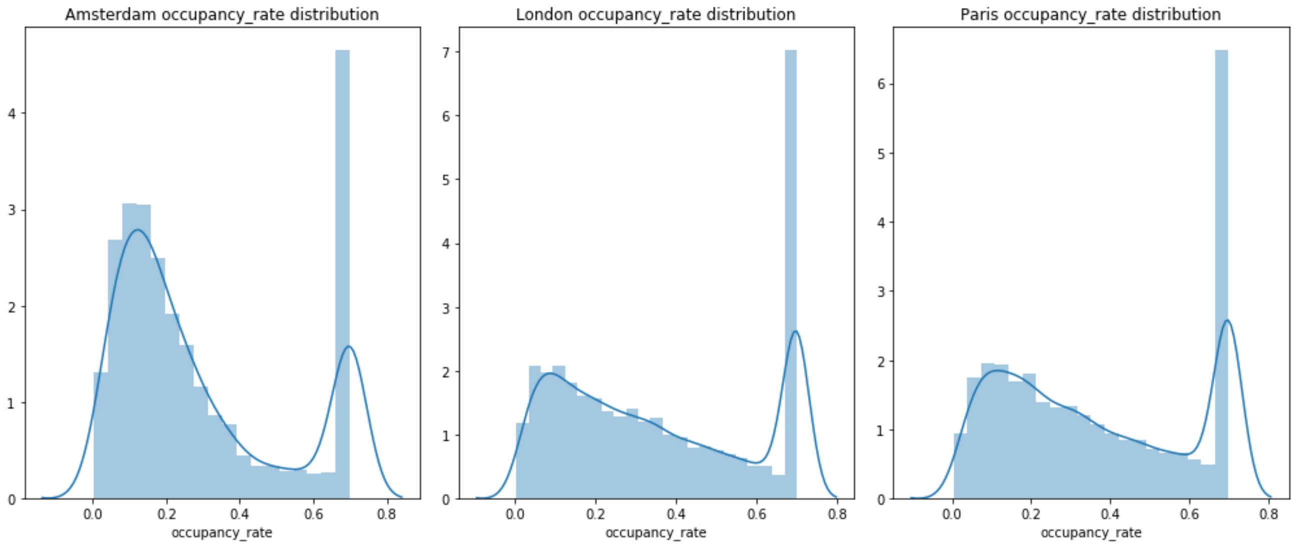


Figure H.1.1. Distribution Histogram with Dependent Variable Occupancy Rate in Amsterdam, Paris and London.

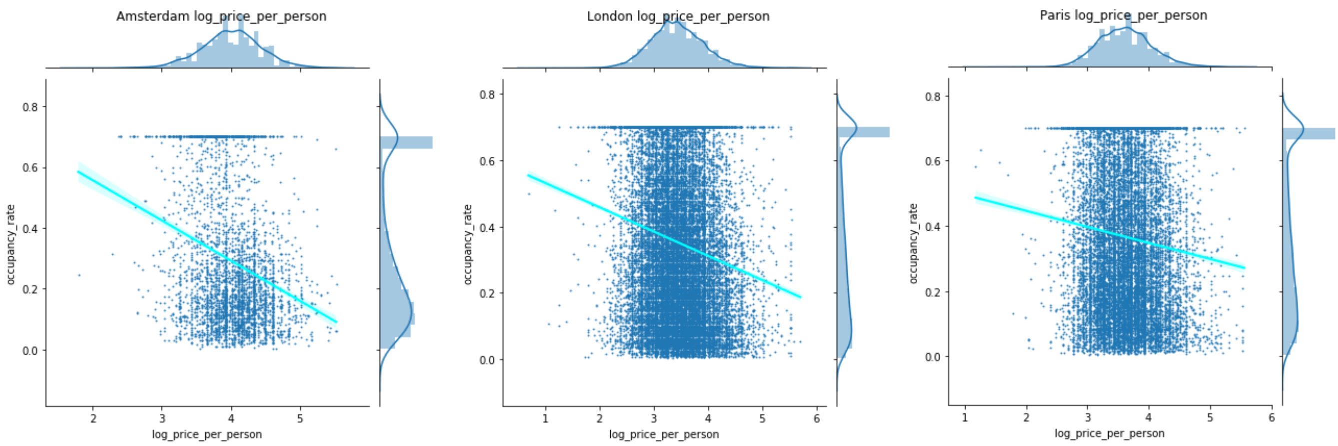


Figure H.1.2. Scatterplot and Distribution Histogram with Dependent Variable Occupancy Rate and the Logarithmic Price per Person in Amsterdam, Paris and London.

Appendix H continued

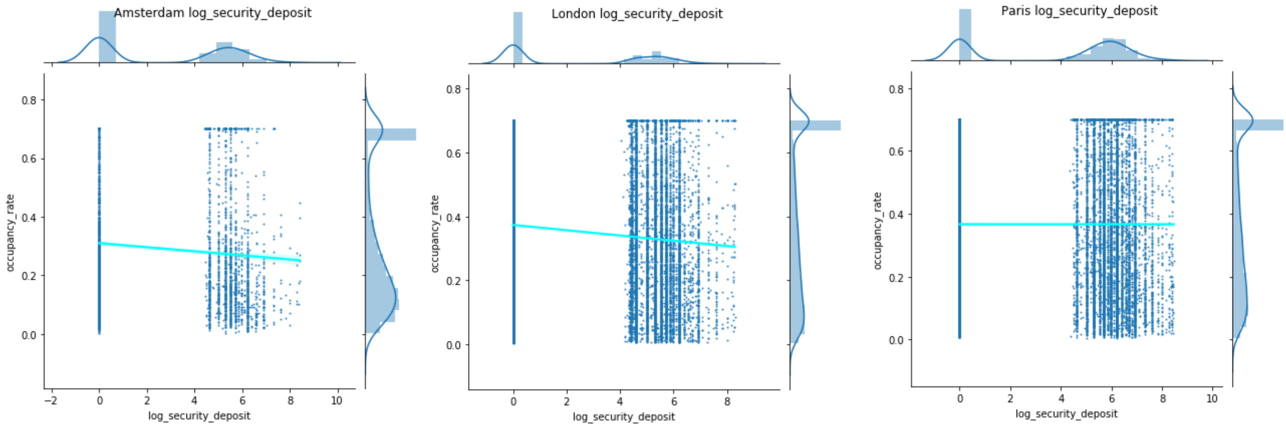


Figure H.1.3. Scatterplot and Distribution Histogram with Dependent Variable Occupancy Rate and the Logarithmic Security Deposit in Amsterdam, Paris and London.

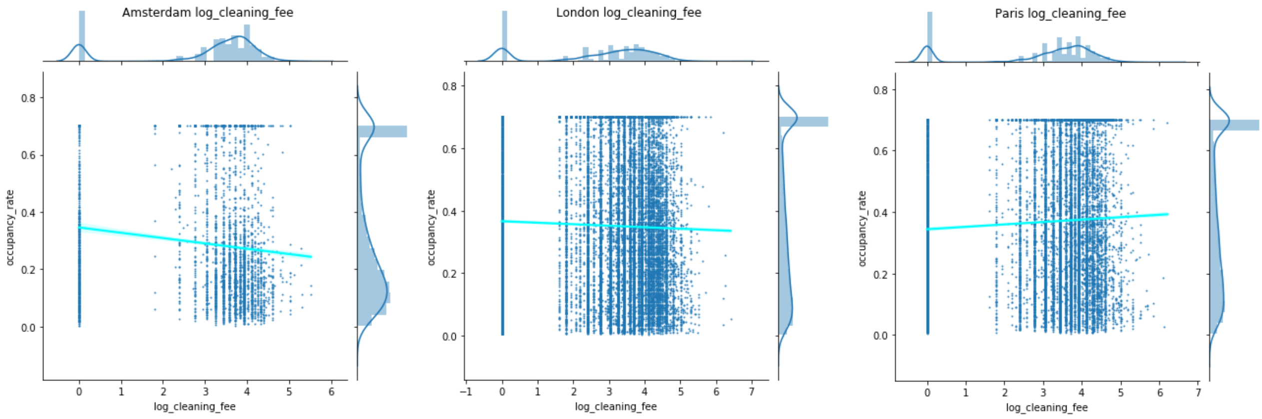


Figure H.1.4. Scatterplot and Distribution Histogram with Dependent Variable Occupancy Rate and the Logarithmic Cleaning Fee in Amsterdam, Paris and London.

Appendix H continued

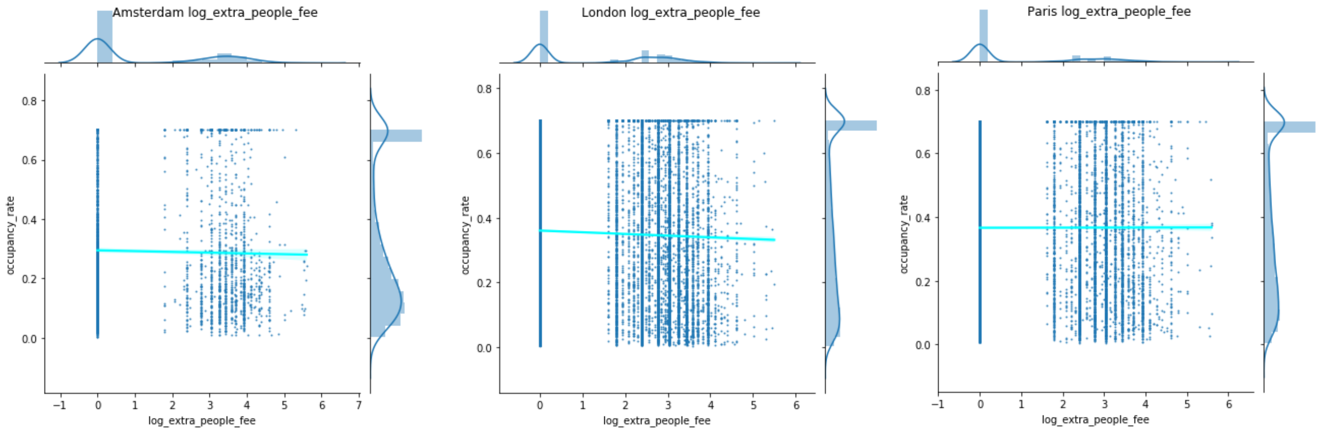


Figure H.1.5. Scatterplot and Distribution Histogram with Dependent Variable Occupancy Rate and the Logarithmic Extra People Fee in Amsterdam, Paris and London.

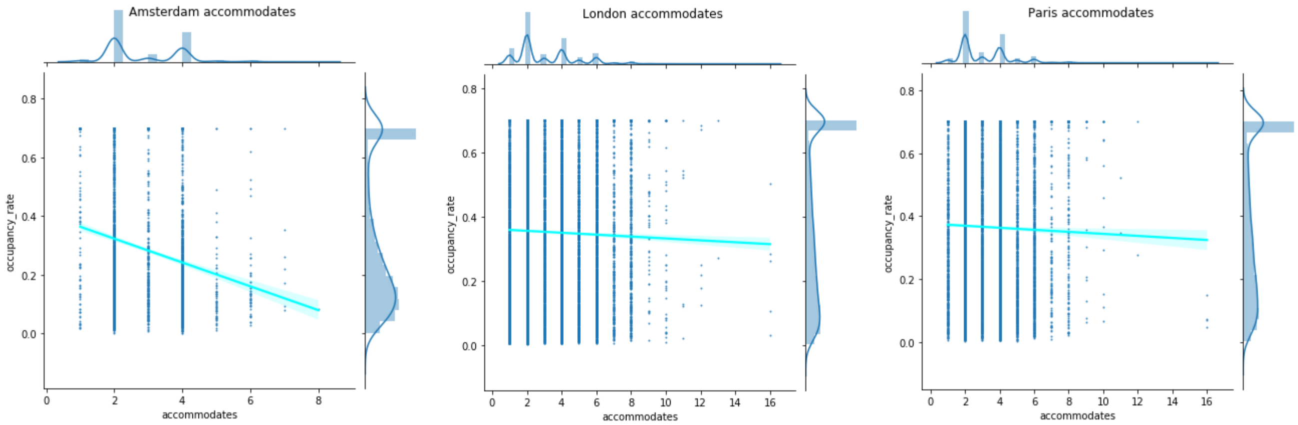


Figure H.1.6. Scatterplot Distribution Histogram with Dependent Variable Occupancy Rate and the Accommodation Capacity in Amsterdam, Paris and London.

Appendix H continued

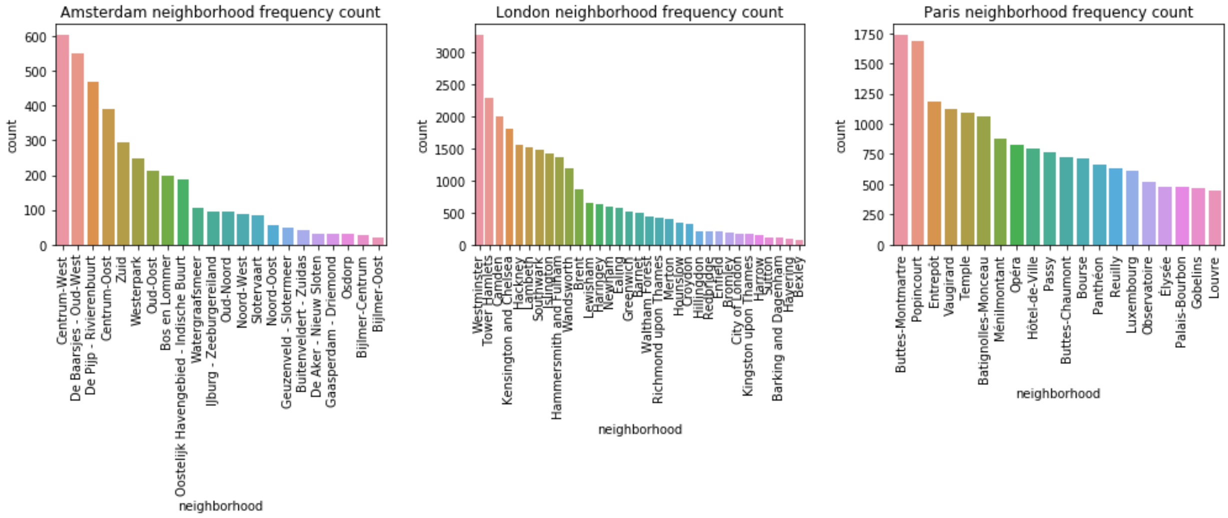


Figure H.2.1. Distribution Frequency Count per Neighborhood in Amsterdam, Paris and London.

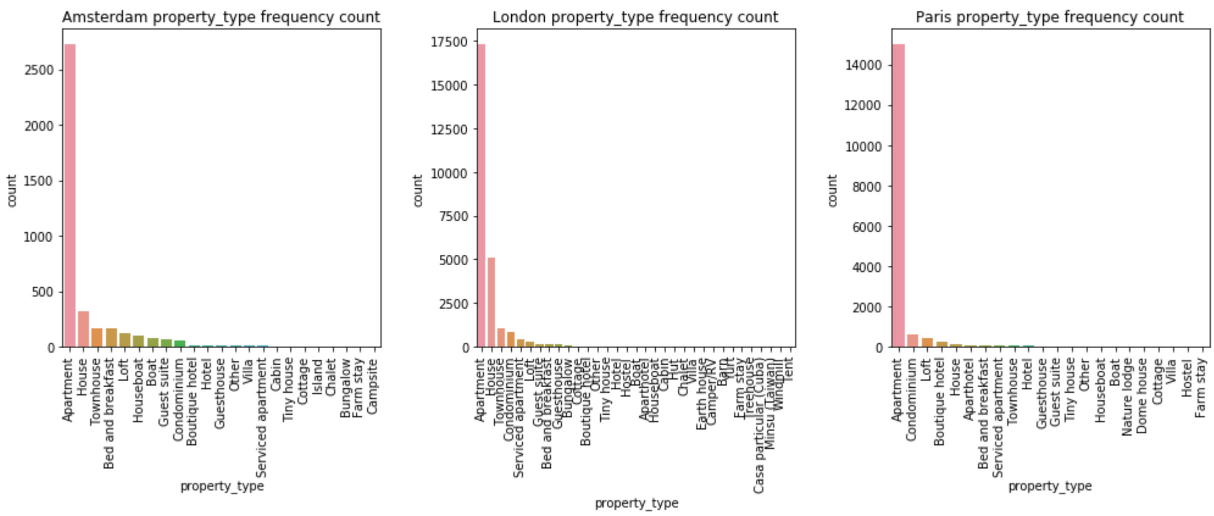


Figure H.2.2. Distribution Frequency Count per Property Type in Amsterdam, Paris and London.

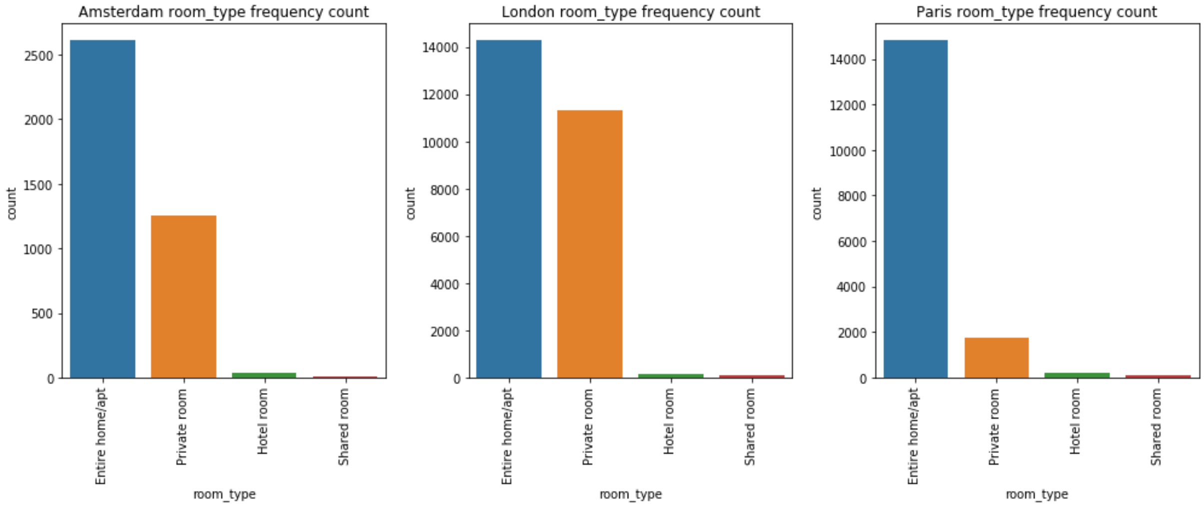


Figure H.2.3. Distribution Frequency Count per Room Type in Amsterdam, Paris and London.

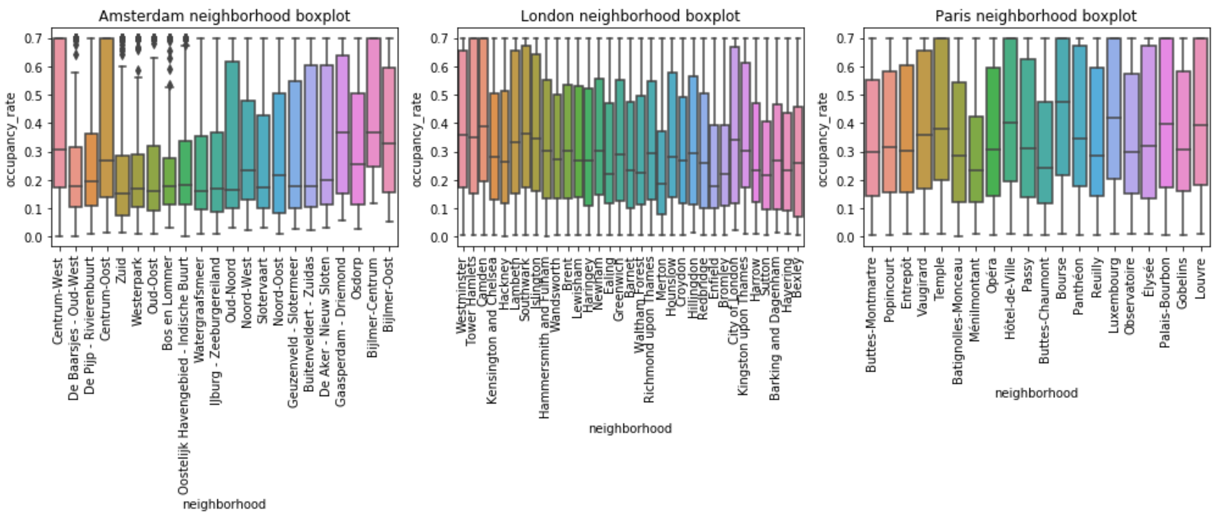


Figure H.2.4. Boxplot with Dependent Variable Occupancy Rate and Neighborhood in Amsterdam, Paris and London.

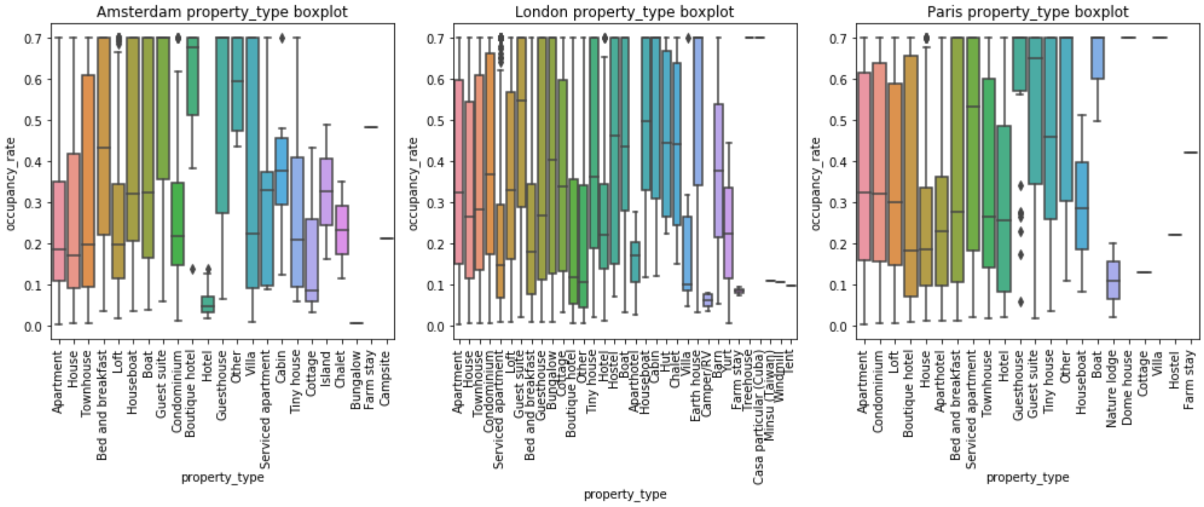


Figure H.2.5. Boxplot with Dependent Variable Occupancy Rate and Property Type in Amsterdam, Paris and London.

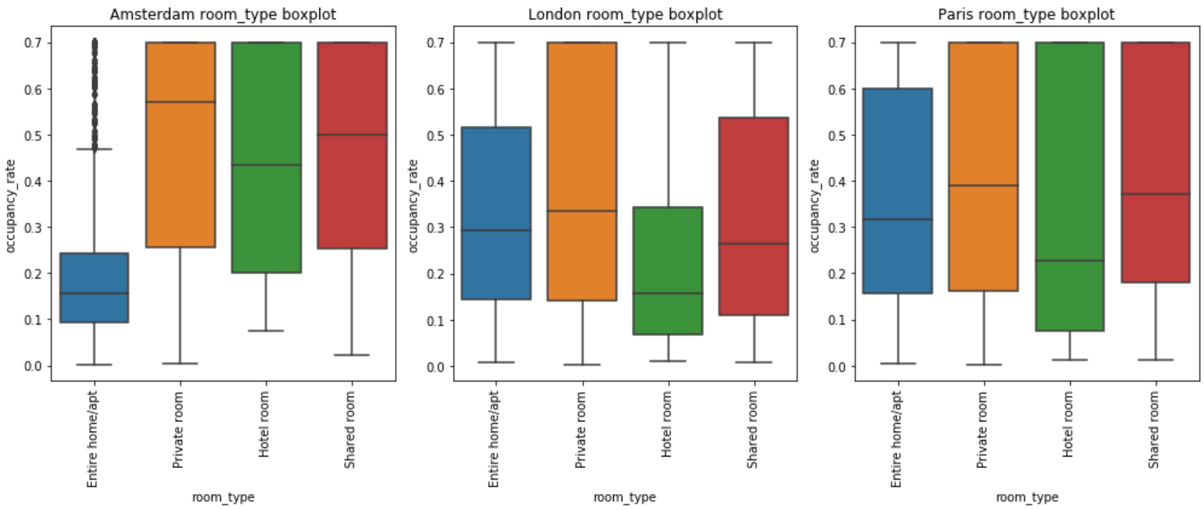


Figure H.2.6. Boxplot with Dependent Variable Occupancy Rate and Room Type in Amsterdam, Paris and London.

Appendix H continued

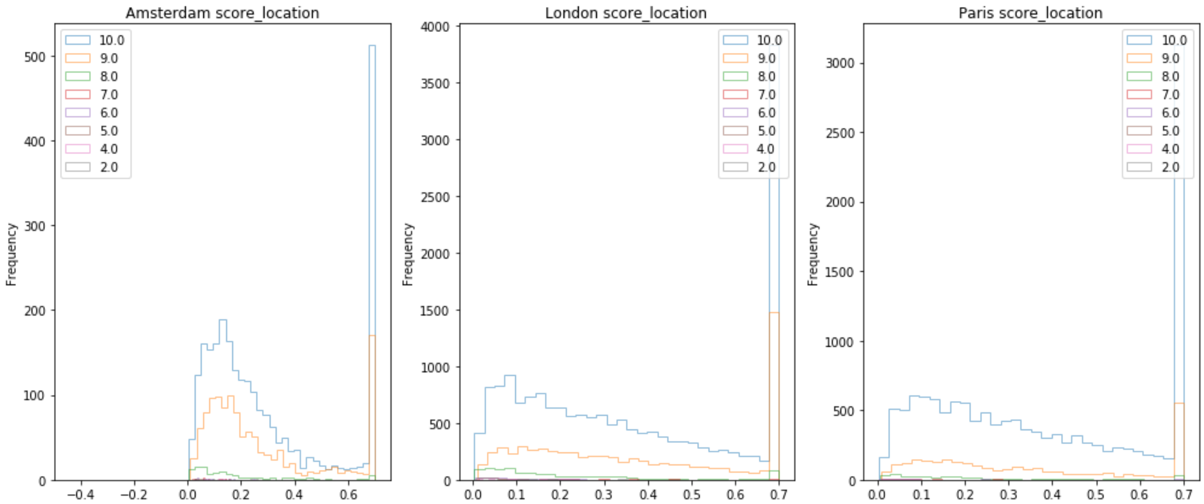


Figure H.3.1. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Location', in Amsterdam, Paris and London.

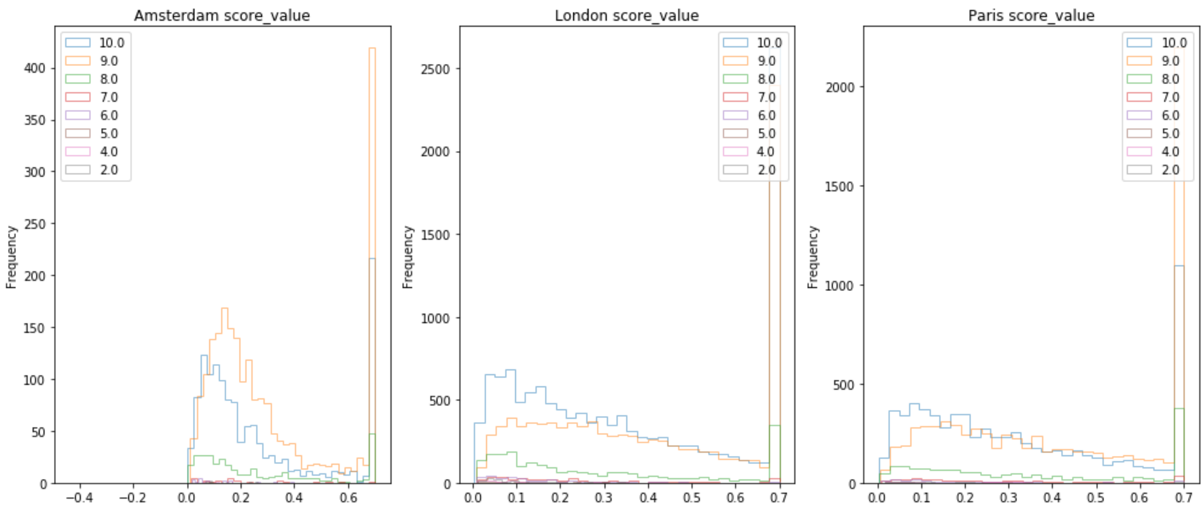


Figure H.3.2. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Value', in Amsterdam, Paris and London.

Appendix H continued

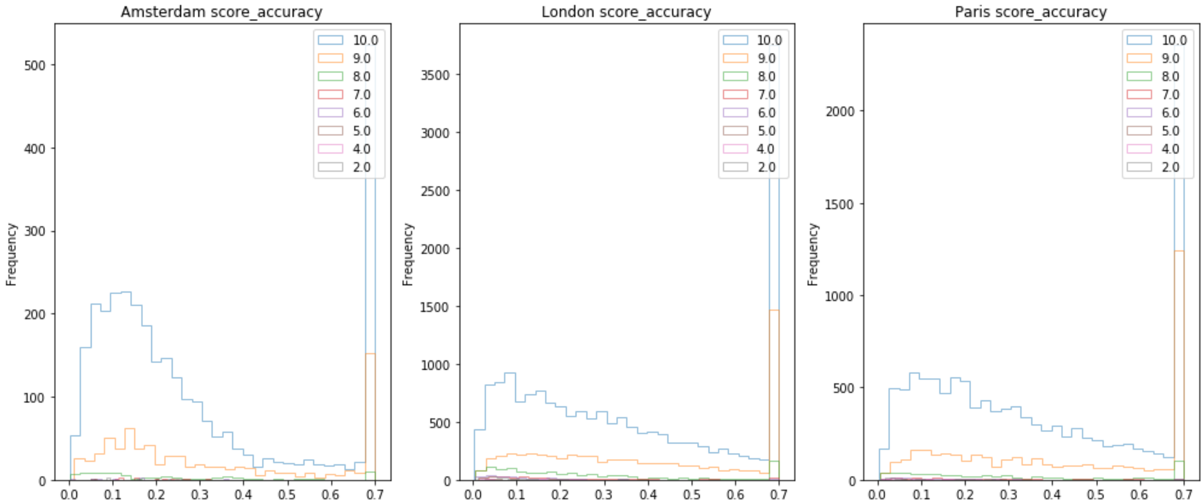


Figure H.3.3. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Accuracy', in Amsterdam, Paris and London.

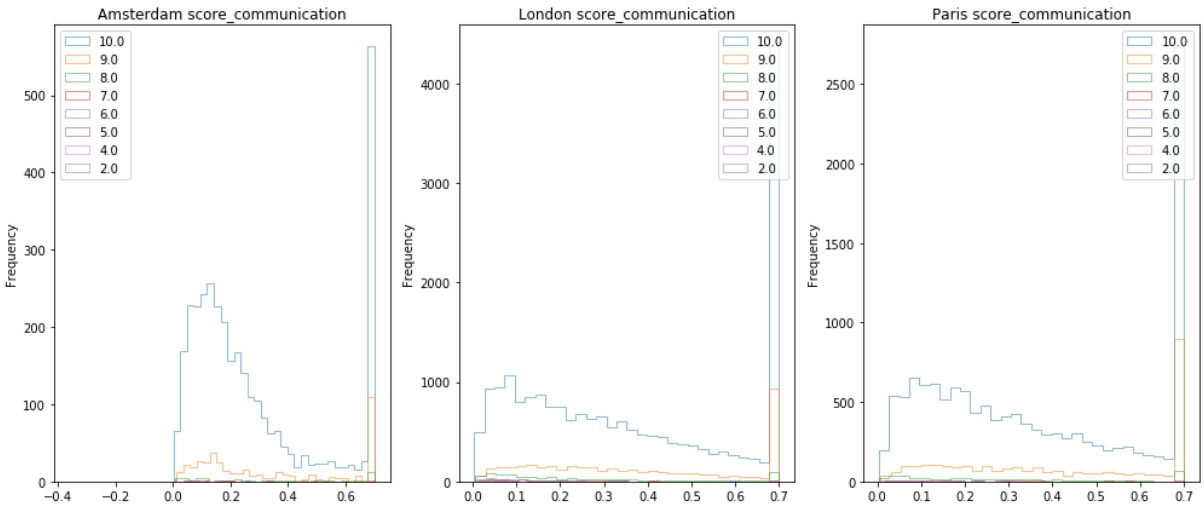


Figure H.3.4. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Communication', in Amsterdam, Paris and London.

Appendix H continued

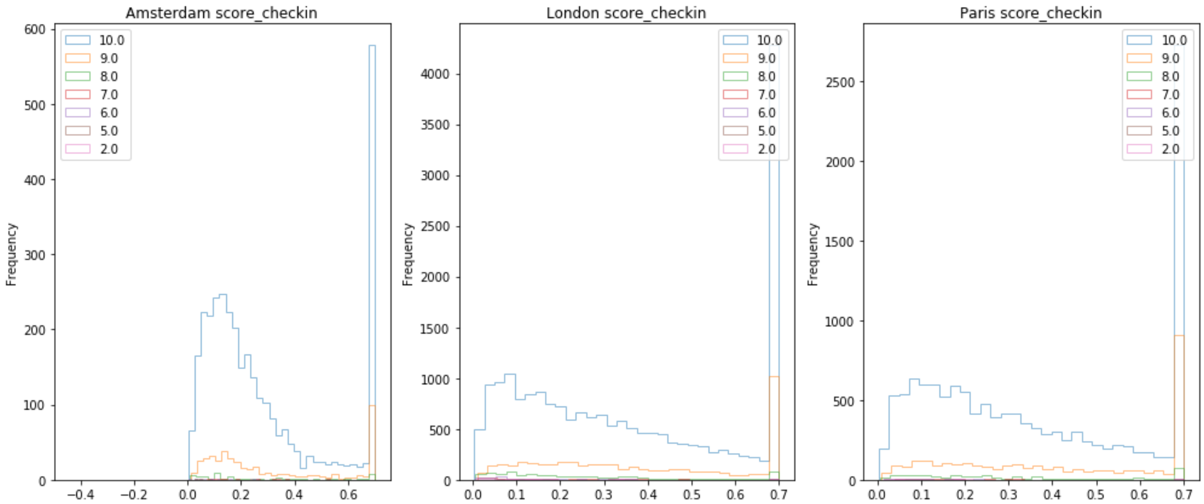


Figure H.3.5. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Check-in', in Amsterdam, Paris and London.

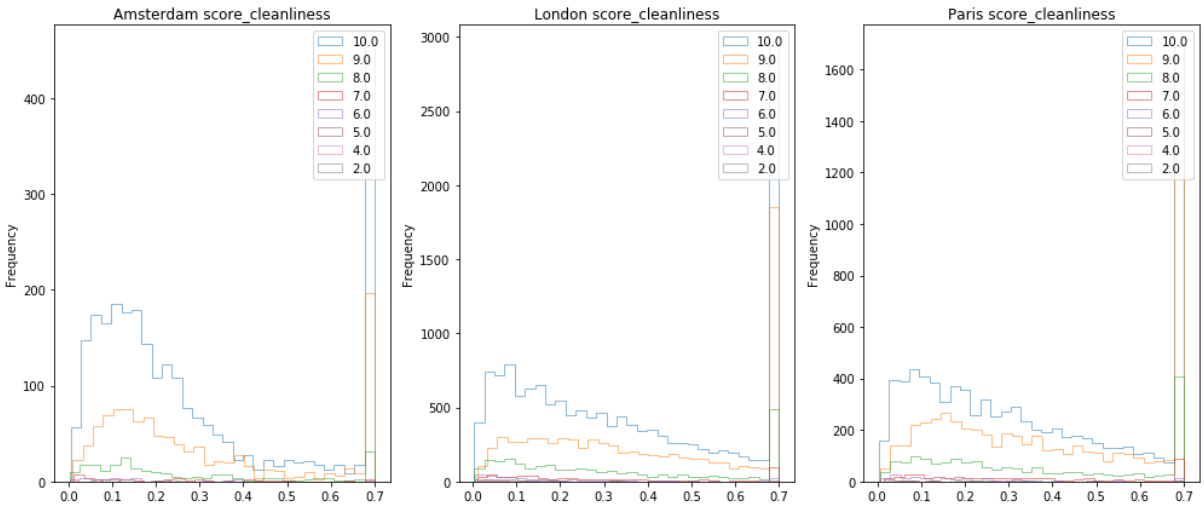


Figure H.3.6. Distribution of Occupancy Rate Conditioned on Open Badge 'Score Cleanliness, in Amsterdam, Paris and London.

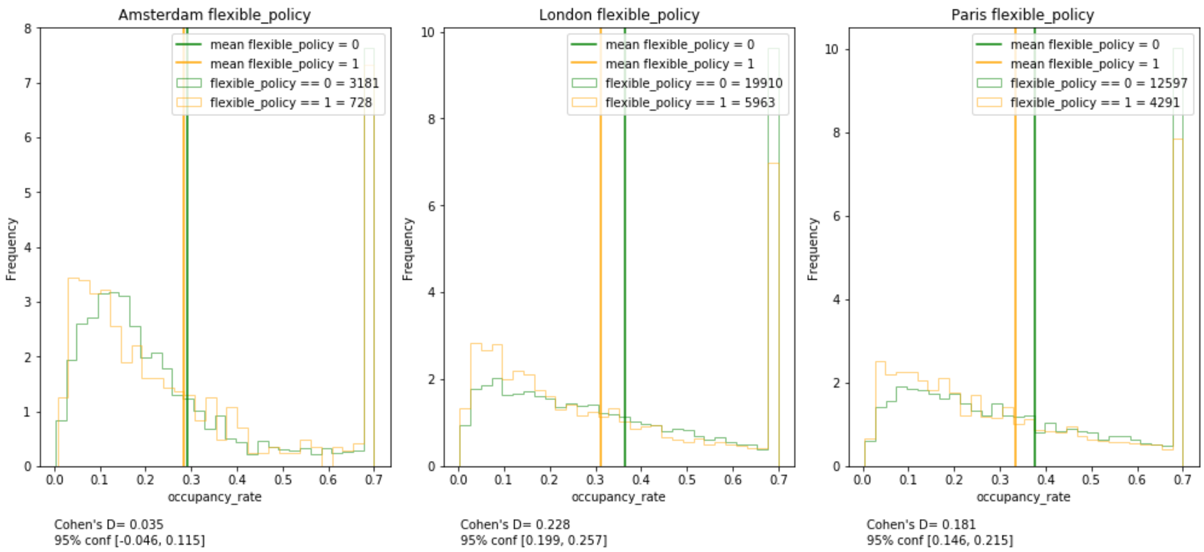


Figure H.4.1. Distribution of Occupancy Rate Conditioned on Flexible Policy Including the Effect Size Estimate, in Amsterdam, Paris and London.

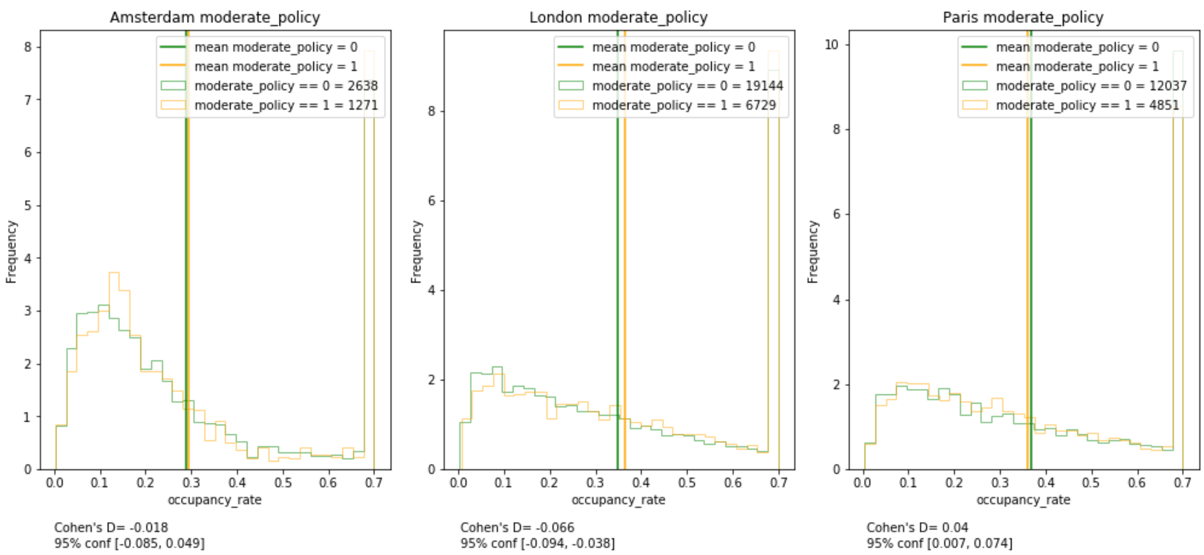


Figure H.4.2. Distribution of Occupancy Rate Conditioned on Moderate Policy Including the Effect Size Estimate, in Amsterdam, Paris and London.

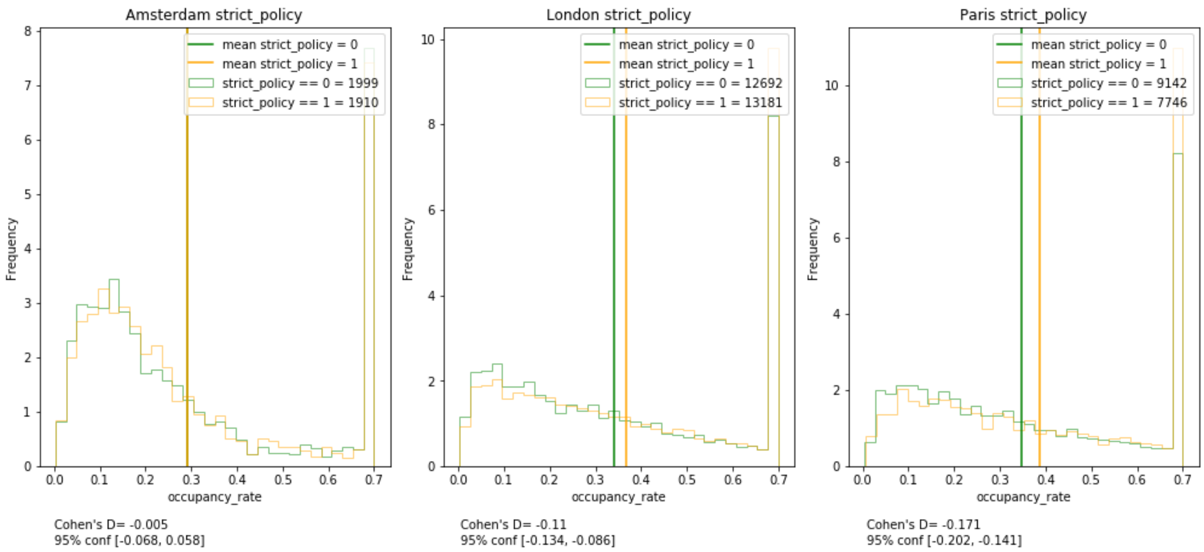


Figure H.4.3. Distribution of Occupancy Rate Conditioned on Strict Policy Including the Effect Size Estimate, in Amsterdam, Paris and London.

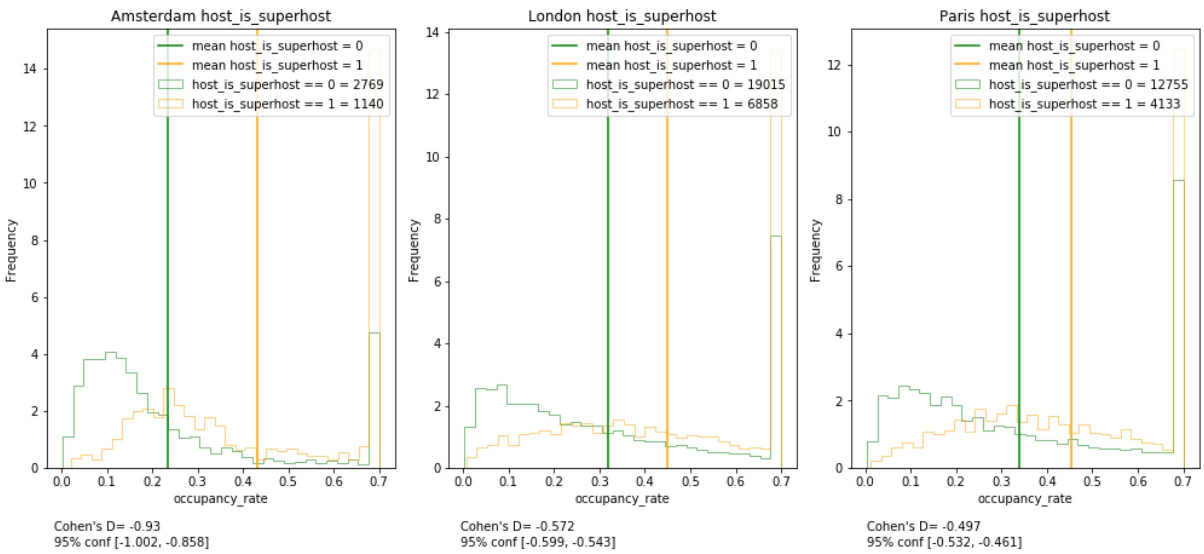


Figure H.4.4. Distribution of Occupancy Rate Conditioned on Open Badge 'Superhost' Including the Effect Size Estimate, in Amsterdam, Paris and London.

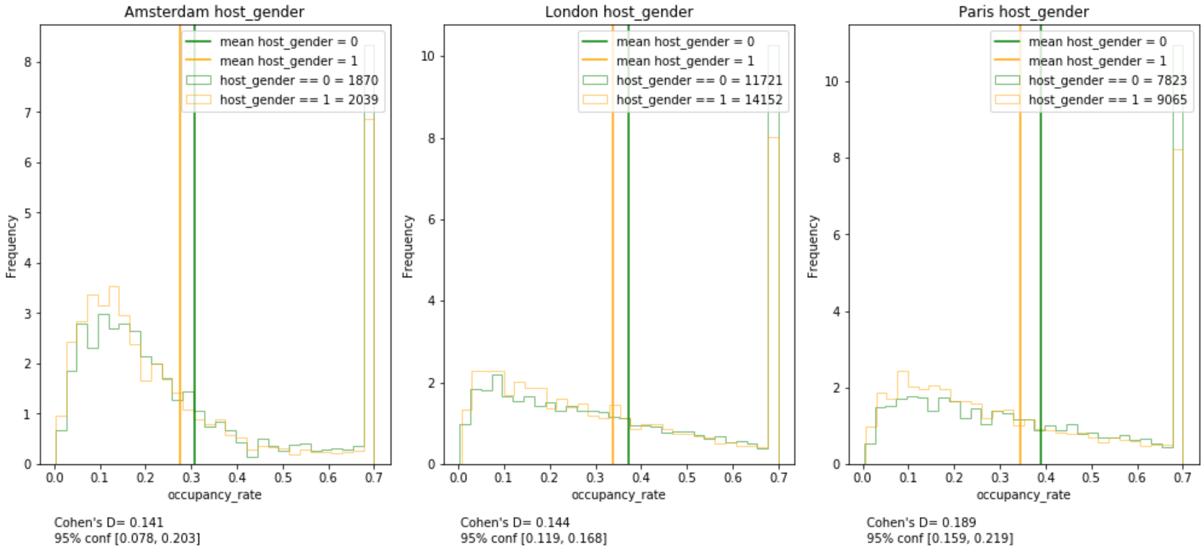


Figure H.4.5. Distribution of Occupancy Rate Conditioned on Open Badge Host Gender Including the Effect Size Estimate, in Amsterdam, Paris and London.

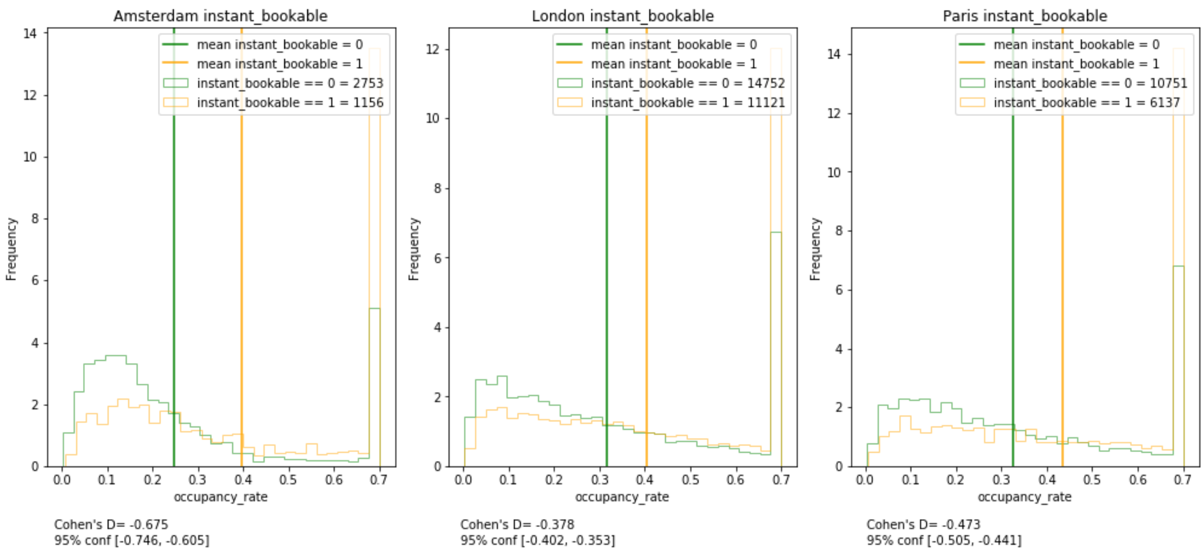


Figure H.4.6. Distribution of Occupancy Rate Conditioned on Control Variable ‘Instant Bookable’ Including the Effect Size Estimate, in Amsterdam, Paris and London.

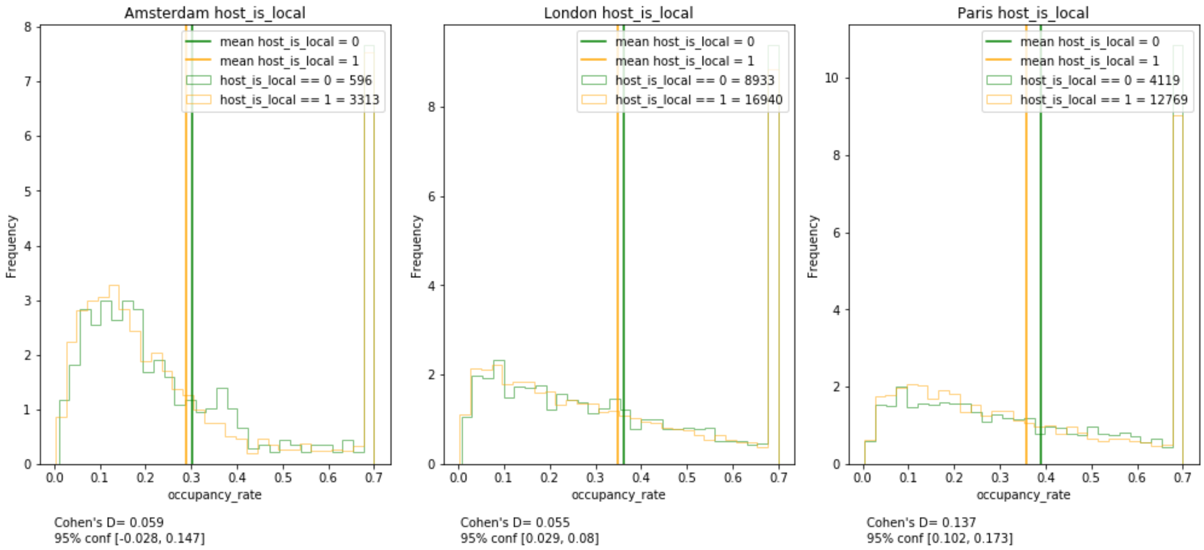


Figure H.4.7. Distribution of Occupancy Rate Conditioned on Control Variable 'Host is Local' Including the Effect Size Estimate, in Amsterdam, Paris and London.

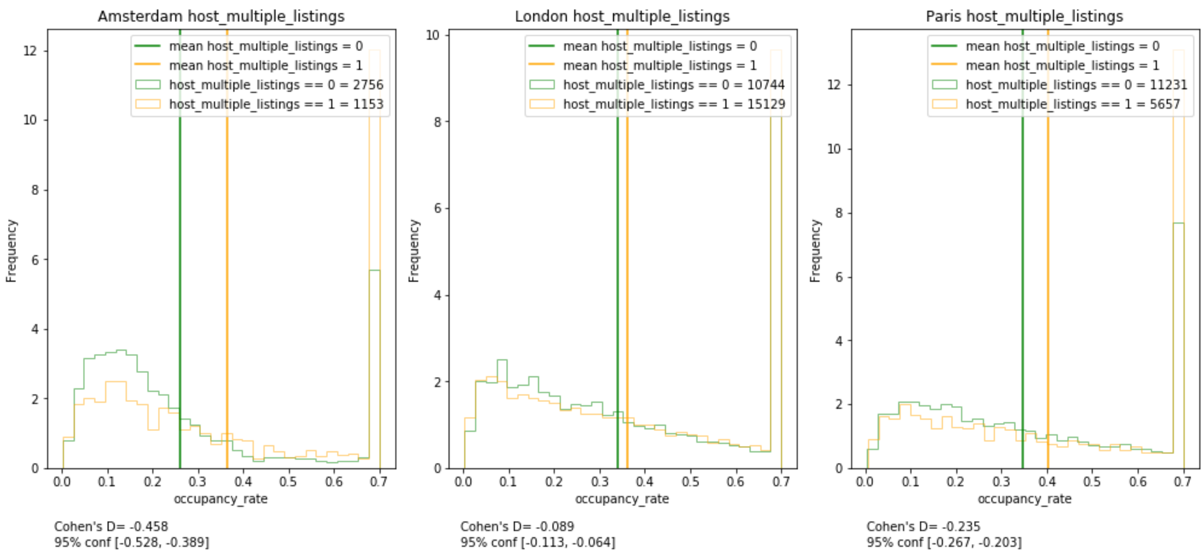


Figure H.4.8. Distribution of Occupancy Rate Conditioned on Control Variable 'Host Multiple Listings' Including the Effect Size Estimate, in Amsterdam, Paris and London.

Appendix H continued

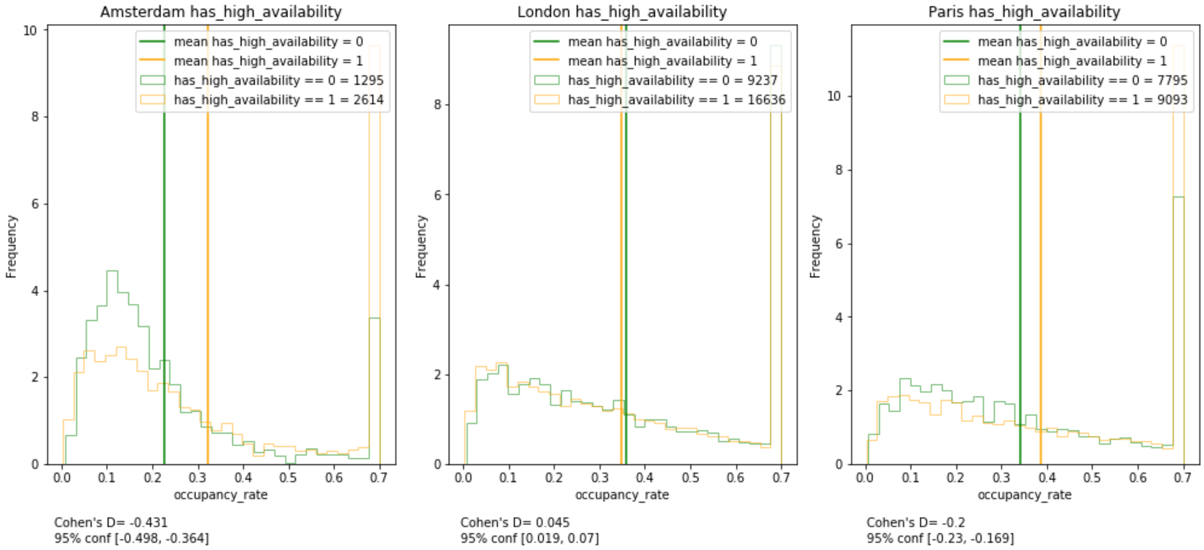


Figure H.4.9. Distribution of Occupancy Rate Conditioned on Control Variable ‘Has High Availability’ Including the Effect Size Estimate, in Amsterdam, Paris and London.

Appendix I

The Regression Model Outcomes for Amsterdam, Paris and London

Table I.1. Extended Regression Results with Dependent Variable Occupancy Rate, Amsterdam.

VARIABLES	(1) Occupancy Rate Model 1	(2) Occupancy Rate Model 2	(3) Occupancy Rate Model 3	(4) Occupancy Rate Model 4
Cancellation Policy				
moderate_policy		0.032***	0.031***	0.026**
		(0.008)	(0.008)	(0.012)
strict_policy		0.019**	0.019**	0.018
		(0.008)	(0.008)	(0.011)
Open Badges				
host_is_superhost		0.126***	0.126***	0.118***
		(0.007)	(0.007)	(0.01)
score_checkin_c9		0.001	0.000	0.001
		(0.009)	(0.009)	(0.012)
score_communication_c9		-0.008	-0.008	0.005
		(0.009)	(0.009)	(0.012)
score_accuracy_c9		0.004	0.004	-0.001
		(0.008)	(0.008)	(0.012)
score_value_c9		-0.009*	-0.009	-0.013
		(0.005)	(0.005)	(0.008)
score_location_c9		0.024***	0.024***	0.021***
		(0.006)	(0.006)	(0.008)
score_cleanliness_c9		0.008	0.009*	0.01
		(0.005)	(0.005)	(0.008)
Moderators				
host_gender			-0.019***	-0.015
			(0.005)	(0.017)
host_gender*moderate_policy				0.013
				(0.016)
host_gender*strict_policy				0.002
				(0.015)
host_gender*host_is_superhost				0.016
				(0.012)
host_gender*score_checkin_c9				-0.001
				(0.017)
host_gender*score_communication_c9				-0.029*
				(0.017)
host_gender*score_accuracy_c9				0.011
				(0.016)
host_gender*score_value_c9				0.008
				(0.011)
host_gender*score_location_c9				0.004
				(0.01)
host_gender*score_cleanliness_c9				-0.002
				(0.010)
Control Variables				
log_price_per_person	-0.114***	-0.125***	-0.126***	-0.127***
	(0.008)	(0.008)	(0.008)	(0.008)
log_security_deposit	0.000	-0.000	1.38e-06	4.48e-05
	(0.001)	(0.001)	(0.001)	(0.001)
log_cleaning_fee	-0.002	-0.003	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
log_extra_people_fee	-0.004**	-0.004**	-0.004**	-0.004**

	(0.002)	(0.002)	(0.002)	(0.002)
log_distance_cost	-0.069***	-0.053***	-0.052***	-0.052***
	(0.006)	(0.006)	(0.006)	(0.006)
instant_bookable	0.072***	0.068***	0.066***	0.066***
	(0.007)	(0.007)	(0.007)	(0.007)
host_is_local	0.018**	0.001	0.001	0.002
	(0.009)	(0.008)	(0.008)	(0.008)
host_multiple_listings	0.005	0.005	0.005	0.006
	(0.007)	(0.007)	(0.007)	(0.007)
has_high_availability	0.050***	0.053***	0.052***	0.052***
	(0.006)	(0.006)	(0.006)	(0.006)
accommodates	-0.014***	-0.014***	-0.015***	-0.014***
	(0.003)	(0.003)	(0.003)	(0.003)
neighborhood_n	-0.002***	-0.003***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
property_type_n	0.001	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
room_type_n	0.106***	0.0874***	0.0871***	0.0867***
	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.608***	0.626***	0.643***	0.645***
	(0.042)	(0.042)	(0.042)	(0.043)
Observations	3,909	3,909	3,909	3,909
R-squared	0.402	0.470	0.471	0.473
Adj. R-squared	0.400	0.467	0.468	0.468

NOTE: Variables 'log_variables' transformed to (natural) logarithms by means of the formula: log (variable+1), with +1 to avoid missing values.

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix I continued

Table I.2. Extended Regression Results with Dependent Variable Occupancy Rate, Paris.

VARIABLES	(1) Occupancy Rate Model 1	(2) Occupancy _Rate Model 2	(3) Occupancy _Rate Model 3	(4) Occupancy _Rate Model 4
Cancellation Policy (H1)				
moderate_policy		0.025***	0.025***	0.023***
		(0.005)	(0.005)	(0.007)
strict_policy		0.046***	0.046***	0.054***
		(0.004)	(0.004)	(0.006)
Open Badges (H2)				
host_is_superhost		0.124***	0.123***	0.121***
		(0.004)	(0.004)	(0.006)
score_checkin_c9		-0.004	-0.004	-0.007
		(0.004)	(0.004)	(0.006)
score_communication_c9		0.003	0.003	0.006
		(0.004)	(0.004)	(0.006)
score_accuracy_c9		0.005	0.005	0.006
		(0.004)	(0.004)	(0.005)
score_value_c9		-0.019***	-0.019***	-0.018***
		(0.003)	(0.003)	(0.005)
score_location_c9		0.049***	0.049***	0.048***
		(0.003)	(0.003)	(0.005)
score_cleanliness_c9		-0.000	0.001	0.001
		(0.003)	(0.003)	(0.004)
Moderators				
host_gender			-0.033***	-0.028***
			(0.003)	(0.009)
host_gender*moderate_policy				0.003
				(0.009)
host_gender*strict_policy				-0.014*
				(0.008)
host_gender*host_is_superhost				0.005
				(0.008)
host_gender*score_checkin_c9				0.006
				(0.008)
host_gender*score_communication_c9				-0.007
				(0.008)
host_gender*score_accuracy_c9				-0.001
				(0.008)
host_gender*score_value_c9				-0.003
				(0.006)
host_gender*score_location_c9				0.0024
				(0.007)
host_gender*score_cleanliness_c9				0.000
				(0.005)
Control Variables				
log_price_per_person	-0.087***	-0.110***	-0.111***	-0.111***
	(0.004)	(0.004)	(0.004)	(0.004)
log_security_deposit	0.001	-0.000	-1.54e-05	1.65e-05
	(0.001)	(0.001)	(0.001)	(0.001)
log_cleaning_fee	0.009***	0.005***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
log_extra_people_fee	-0.001	-0.005***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
log_distance_cost	-0.062***	-0.052***	-0.052***	-0.052***

	(0.005)	(0.005)	(0.005)	(0.005)
instant_bookable	0.100***	0.099***	0.097***	0.097***
	(0.004)	(0.004)	(0.004)	(0.004)
host_is_local	-0.015***	-0.021***	-0.021***	-0.021***
	(0.004)	(0.004)	(0.004)	(0.004)
host_multiple_listings	0.035***	0.032***	0.030***	0.03***
	(0.004)	(0.004)	(0.004)	(0.004)
has_high_availability	0.042***	0.041***	0.040***	0.040***
	(0.004)	(0.004)	(0.005)	(0.005)
accommodates	-0.015***	-0.015***	-0.016***	-0.016***
	(0.001)	(0.001)	(0.001)	(0.001)
neighborhood_n	0.002***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
property_type_n	-9.34e-05	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
room_type_n	0.012***	0.01**	0.008***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Constant	0.618***	0.643***	0.665***	0.664***
	(0.017)	(0.018)	(0.018)	(0.018)
Observations	16,888	16,888	16,888	16,888
R-squared	0.092	0.162	0.166	0.167
Adj. R-squared	0.092	0.161	0.165	0.165

NOTE: Variables ‘log_variables’ transformed to (natural) logarithms by means of the formula: log(variable+1), with +1 to avoid missing values.

Robust standard errors in parentheses

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

Appendix I continued

Table I.3. Extended Regression Results with Dependent Variable Occupancy Rate, London.

VARIABLES	(1) Occupancy Rate Model 1	(2) Occupancy Rate Model 2	(3) Occupancy Rate Model 3	(4) Occupancy Rate Model 4
Cancellation Policy (H1)				
moderate_policy		0.048***	0.049***	0.059***
		(0.004)	(0.004)	(0.006)
strict_policy		0.051***	0.051***	0.066***
		(0.004)	(0.004)	(0.005)
Open Badges (H2)				
host_is_superhost		0.136***	0.137***	0.128***
		(0.003)	(0.003)	(0.005)
score_checkin_c9		0.000	0.000	-0.001
		(0.002)	(0.002)	(0.003)
score_communication_c9		0.009***	0.009***	0.013***
		(0.002)	(0.002)	(0.003)
score_accuracy_c9		0.012***	0.012***	0.013***
		(0.002)	(0.002)	(0.003)
score_value_c9		0.001	0.001	-0.001
		(0.002)	(0.002)	(0.003)
score_location_c9		0.017***	0.017***	0.016***
		(0.002)	(0.002)	(0.003)
score_cleanliness_c9		-0.002	-0.001	-0.001
		(0.002)	(0.002)	(0.003)
Moderators				
host_gender				-0.009
				(0.006)
host_gender*moderate_policy				-0.019**
				(0.008)
host_gender*strict_policy				-0.026***
				(0.007)
host_gender*host_is_superhost				0.015**
				(0.006)
host_gender*score_checkin_c9				0.0022
				(0.005)
host_gender*score_communication_c9				-0.007
				(0.005)
host_gender*score_accuracy_c9				-0.008
				(0.005)
host_gender*score_value_c9				0.002
				(0.005)
host_gender*score_location_c9				0.003
				(0.004)
host_gender*score_cleanliness_c9				-0.001
				(0.004)
Control Variables				
log_price_per_person	-0.107***	-0.118***	-0.117***	-0.117***
	(0.003)	(0.003)	(0.003)	(0.003)
log_security_deposit	-0.004***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
log_cleaning_fee	0.001	-0.002**	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
log_extra_people_fee	-0.007***	-0.01***	-0.01***	-0.01***
	(0.001)	(0.001)	(0.001)	(0.001)
log_distance_cost	-0.064***	-0.064***	-0.062***	-0.062***

	(0.002)	(0.002)	(0.002)	(0.002)
instant_bookable	0.071***	0.087***	0.085***	0.085***
	(0.003)	(0.003)	(0.003)	(0.003)
host_is_local	-0.010***	-0.021***	-0.021***	-0.021***
	(0.003)	(0.003)	(0.003)	(0.003)
host_multiple_listings	-0.001	0.004	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)
has_high_availability	-0.004	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
accommodates	-0.009***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
neighborhood_n	4.28e-05	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
property_type_n	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
room_type_n	0.005**	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.822***	0.779***	0.789***	0.780***
	(0.017)	(0.017)	(0.017)	(0.017)
Observations	25,873	25,873	25,873	25,873
R-squared	0.108	0.205	0.207	0.208
Adj. R-squared	0.108	0.204	0.206	0.207

NOTE: Variables ‘log_variables’ transformed to (natural) logarithms by means of the formula: log (variable+1), with +1 to avoid missing values.

Robust standard errors in parentheses

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

References

Achinstein, B., & Ogawa, R. T. (2012). New teachers of color and culturally responsive teaching in an era of educational accountability: Caught in a double bind. *Journal of educational change*, *13*(1), 1-39.

Airbnb (2019). *Women Hosts Are Leading the Way on Airbnb*. Retrieved from https://news.airbnb.com/women-hosts-are-leading-the-way-on-airbnb/

Airbnb (n.d.). How to start hosting? Available at: from https://www.airbnb.com/d/setup (Accessed March 10, 2020)

Airbnb (n.d.). *Rent out your house, apartment or room on Airbnb*. Available at: https://www.airbnb.com/host/homes (Accessed March 7, 2020).

Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological science*, *7*(3), 136-141.

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of personality and Social Psychology*, *65*(2), 272.

Benkler, Y. (2004). Sharing nicely: On shareable goods and the emergence of sharing as a modality of economic production. *Yale LJ*, *114*, 273.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, *74*(2), 170.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review*, *94*(4), 991-1013.

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of public economics*, *145*, 27-41.

Brucks, M., Zeithaml, V. A., & Naylor, G. (2000). Price and brand name as indicators of quality dimensions for consumer durables. *Journal of the academy of marketing science*, *28*(3), 359-374.

Buchanan, L., Bui, Q., & Patel, J. K. (2020). Black Lives Matter May Be the Largest Movement in U.S. History. Retrieved from https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html

Cansoy, M., & Schor, J. B. (2017). Who gets to share in the sharing economy: Racial discrimination on Airbnb. *Working paper*. Boston College.

Catalyst, inc. (2007). *The double-bind dilemma for women in leadership: Damned if you do, doomed if you don't*. Catalyst.

Chong, A., & Calderon, C. (2000). Causality and feedback between institutional measures and economic growth. *Economics & Politics*, *12*(1), 69-81.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of management*, *37*(1), 39-67.

Craney, T. A., & Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, *14*(3), 391-403.

Curry, D. J., & Faulds, D. J. (1986). Indexing product quality: Issues, theory, and results. *Journal of Consumer Research*, *13*(1), 134-145.

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, *7*(4), 145-147.

Dellaert, B. G. (2019). The consumer production journey: marketing to consumers as co-producers in the sharing economy. *Journal of the Academy of Marketing Science*, 47(2), 238-254.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological review*, *109*(3), 573.

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of Airbnb.com. *Harvard Business School NOM Unit Working Paper*, (14-054).

Edelman, B., Luca, M., and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics, 9*(2), 1-22.

Elsesser, K. M., & Lever, J. (2011). Does gender bias against female leaders persist? Quantitative and qualitative data from a large-scale survey. *Human Relations*, *64*(12), 1555-1578.

Farronato, C., & Fradkin, A. (2018). *The welfare effects of peer entry in the accommodation market: The case of airbnb* (No. w24361). National Bureau of Economic Research.

Ferber, M. A., & Huber, J. A. (1975). Sex of student and instructor: A study of student bias. *American Journal of Sociology*, *80*(4), 949-963.

Fine, C. (2010). *Delusions of gender: How our minds, society, and neurosexism create difference*. WW Norton & Company.

Fiske, A. P., Haslam, N., & Fiske, S. T. (1991). Confusing one person with another: what errors reveal about the elementary forms of social relations. *Journal of personality and social psychology*, *60*(5), 656.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: nature, purpose, and findings. *Journal of marketing*, *60*(4), 7-18.

Fraiberger, S. P., & Sundararajan, A. (2015). Peer-to-peer rental markets in the sharing economy. *NYU Stern School of Business research paper*, *6*.

Frenken, K., & Schor, J. (2019). Putting the sharing economy into perspective. In *A Research Agenda for Sustainable Consumption Governance*. Edward Elgar Publishing.

Frenken, K., Meelen, T., Arets, M., Van de Glind, P. (2015. *Smarter Regulation for the Sharing Economy*. The Guardian. https://www.theguardian.com/science/political-science/2015/may/20/smarter-regulation-for-the-sharing-economy, Retrieved April 04, 2020.).

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*, *141*(1), 2.

Ge, Y., Knittel, C. R., MacKenzie, D., & Zoepf, S. (2016). *Racial and gender discrimination in transportation network companies* (No. w22776). National Bureau of Economic Research.

Gourieroux, C., Holly, A., & Monfort, A. (1982). Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. *Econometrica: journal of the Econometric Society*, 63-80.

Guo, L. (2009). Service cancellation and competitive refund policy. *Marketing Science*, *28*(5), 901-917.

Hjorth-Andersen, C. (1984). The concept of quality and the efficiency of markets for consumer products. *Journal of Consumer Research*, *11*(2), 708-718.

Huang, M. H., & Dev, C. S. (2019). Growing the service brand. *International Journal of Research in Marketing*.

iPropertyManagement (n.d.). *Airbnb Statistics [2020]: User & Market Growth Data*. Retrieved from https://ipropertymanagement.com/research/airbnb-statistics

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.

Jiang, P., & Rosenbloom, B. (2005). Customer intention to return online: price perception, attribute-level performance, and satisfaction unfolding over time. *European Journal of Marketing*.

Jovanovic, J., & Devedzic, V. (2015). Open badges: Novel means to motivate, scaffold and recognize learning. *Technology, Knowledge and Learning*, *20*(1), 115-122.

Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American economic review*, *75*(3), 424-440.

Lawrie, E. (2016). *Richard Branson: Our obsession with ownership is at a tipping point*. Retrieved from https://www.thisismoney.co.uk/money/smallbusiness/article-3663869/Richard-Branson-obsession-ownership-tipping-point.html

Lewis, G. (2011). Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors. *American Economic Review*, 101(4): 1535-46.

Li, H., & Srinivasan, K. (2019). Competitive dynamics in the sharing economy: an analysis in the context of Airbnb and hotels. *Marketing Science*, *38*(3), 365-391.

Liu, C., & Arnett, K. P. (2000). Exploring the factors associated with Web site success in the context of electronic commerce. *Information & management*, *38*(1), 23-33.

McClay, C. (2020). Goya Foods: Hispanic brand faces boycott for praising Trump. Retrieved from https://www.bbc.com/news/world-us-canada-53371392

McMeel, G. (2007). *The construction of contracts: interpretation, implication, and rectification*. Oxford University Press.

Mendes, K., Ringrose, J., & Keller, J. (2018). # MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236-246.

Monroe, K. B., & Krishnan, R. (1985). The effect of price on subjective product evaluations. *Perceived quality*, *1*(1), 209-232.

Oghazi, P., Karlsson, S., Hellström, D., & Hjort, K. (2018). Online purchase return policy leniency and purchase decision: Mediating role of consumer trust. *Journal of Retailing and Consumer Services*, *41*, 190-200.

Pansari, A., & Kumar, V. (2017). Customer engagement: the construct, antecedents, and consequences. *Journal of the Academy of Marketing Science*, 45(3), 294-311.

Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard business review*, *86*(1), 25-40.

Power, R. (2020). How 5 Businesses Are Adapting To Life In A Pandemic. Retrieved from https://www.forbes.com/sites/rhettpower/2020/04/26/how-5-businesses-are-adapting-to-life-in-a-pandemic/#37107dc44f5a

Priyadarsini, R., Xuchao, W., & Eang, L. S. (2009). A study on energy performance of hotel buildings in Singapore. *Energy and buildings*, *41*(12), 1319-1324.

Rauch, D., & Schleicher, D. (2015). Like Uber, but for local governmental policy: the future of local regulation of the 'sharing economy'. *George Mason Law & Economics Research Paper*, (15-01).

Ross, S. A. (1973). The economic theory of agency: The principal's problem. *The American economic review*, *63*(2), 134-139.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *science*, *311*(5762), 854-856.

Schor, J. (2014), 'Debating the sharing economy', October, accessed October 23, 2016 at https://www.greattransition.org/publication/debating-the-sharing-economy.

Schor, J. B., Fitzmaurice, C., Carfagna, L. B., Attwood-Charles, W., & Poteat, E. D. (2016). Paradoxes of openness and distinction in the sharing economy. *Poetics*, *54*, 66-81.

Simonsohn, U. (2010). eBay's crowded evenings: Competition neglect in market entry decisions. *Management science*, *56*(7), 1060-1073.

Skaperdas, S. (1991). Conflict and attitudes toward risk. *The American Economic Review*, *81*(2), 116-120.

Spence, M. (1978). Job market signaling. In *Uncertainty in economics* (pp. 281-306). Academic Press.

Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of marketing research*, *50*(2), 277-288.

Srivastava, J., & Lurie, N. (2001). A consumer perspective on price-matching refund policies: Effect on price perceptions and search behavior. *Journal of Consumer Research*, *28*(2), 296-307.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, *62*(2), 207.

Tiago, B., & Veríssimo, C. (2014). Digital marketing and social media: Why bother?. *Business horizons*, *57*(6), 703-708.

van Knippenberg, A., van Twuyver, M., & Pepels, J. (1994). Factors affecting social categorization processes in memory. *British Journal of Social Psychology*, *33*(4), 419-431.

Varadarajan, P. R., & Yadav, M. S. (2002). Marketing strategy and the Internet: an organizing framework. *Journal of the Academy of Marketing Science*, *30*(4), 296-312.

Voss, G. B., Parasuraman, A., & Grewal, D. (1998). The roles of price, performance, and expectations in determining satisfaction in service exchanges. *Journal of marketing*, *62*(4), 46-61.

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management*, *62*, 120-131.

Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer research*, *27*(3), 382-387.

Xie, J., & Gerstner, E. (2007). Service escape: Profiting from customer cancellations. *Marketing Science*, *26*(1), 18-30.

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, *43*, 1-12.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of marketing*, *52*(3), 2-22.

Zeithaml, V. A. (2000). Service quality, profitability, and the economic worth of customers: what we know and what we need to learn. *Journal of the academy of marketing science*, 28(1), 67-85.

Zervas, G., Proserpio, D., & Byers, J. (2015). A first look at online reputation on Airbnb, where every stay is above average. *Where Every Stay is Above Average*.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of marketing research*, 54(5), 687-705.