What is the value of being a superhost?

Aleksander Berentsen, Mariana Rojas Breu, Christopher Waller

What is the Value of Being a Superhost?*

Aleksander Berentsen
University of Basel and FRB St. Louis

Mariana Rojas-Breu
University of Paris Dauphine

Christopher Waller
FRB St. Louis and Deakin University

Abstract

We construct a search model where sellers post prices and produce goods of unknown quality. A match between a buyer and a seller reveals the quality of the seller. We look at the pricing decisions of the sellers in this environment. We then introduce a rating system whereby buyers reveal the seller's type by giving them a 'star' if they are a high quality seller. We show that new sellers charge a low price to attract buyers and if they receive a star they post a high price. Furthermore, high quality sellers sell with a higher probability than new sellers. We show that welfare is higher with a ratings system. Using data on Airbnb rentals to compare the pricing decisions of Superhosts (elite rentals) to non-Superhosts we show that Superhosts: 1) charge higher prices, 2) have a higher occupancy rate and 3) higher revenue than non-Super hosts.

*The views in this paper are those of the authors and do not represent those of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the FOMC. Financial support from the Risk Foundation is gratefully acknowledged.

1 Introduction

For a variety of products, buyers search for sellers of their desired goods. A critical part of the search process is acquiring information about prices. A simple example is shopping on the internet. Sometimes a buyer searches across individual sellers observing a single posted price at each website. Alternatively, the buyer can go to an intermediary site where prices from a variety of sellers are displayed. Examples include Amazon, Yelp, Flixster, Travelocity and Airbnb.

Often the products on these sites are experience goods – the quality can only be ascertained by consuming them. As a result, buyers would like to have an idea of the quality of the products offered to guide their decision making. Consequently, buyers would benefit from observing some indicator of the value of the product. One way to do this is to have buyers give the seller a 'rating' about the quality of the product. If the quality is good, the buyer gives the seller a 'star,' which acts as a signal to future buyers that the seller provides a high quality products. This is a common practice on the sites mentioned above – products get customer reviews on Amazon, restaurants, hotels and movies receive customer satisfaction ratings on Yelp, Trip Advisor and Flixster etc. Airbnb goes a step further by separating an elite group of rentals from all the rest by giving them a 'Superhost' designation. These ratings help the buyer find better quality products.

But what is the value of this rating for sellers? If a seller knows the quality of its product, it can find other means to signal the quality and it does not have to rely on previous buyers to do so. However, buyers are more likely to trust previous buyers and find any signal from the seller to be 'cheap talk'. It may also be the case that the sellers themselves may not know the true quality of their product. For example, new restaurants may not really know how good they are. The same applies for new rental properties on Airbnb. Thus, receiving a positive rating from buyers is good for both the buyers and the sellers.

The problem with ratings is how to get one in the first place. On sites where ratings matter to prospective buyers, not having a rating most likely hurts seller – why go to a seller with no rating as opposed to a seller with a high rating? So new sellers have to provide incentives to attract buyers and get a rating. One way to do this is offer a low price to start – this will attract customers and if the quality is good, the seller receives a high rating. But once the seller receives the rating what is value? It seems clear that the rating allows one to post a higher price for the good or attract even more customers. Our objective in this paper is to model a trading environment where ratings act as a signal of quality and study the dynamic pricing decisions of firms in such an environment.

The theoretical model is a search model where new firms do not know the quality of their good. Buyers prefer higher quality to lower quality goods. Quality is only revealed once a match occurs. We show that in a model without ratings, firms are in a pooling equilibrium and set the same price regardless of quality. We then introduce a rating system in which firms that are revealed as high quality producers receive a 'star' indicating their type. We show that new sellers will charge a low price (relative to the pooling price) in order to attract buyers. If a star is received, they then charge a higher price then the pooling price. Thus, a rating system generates a non-degenerate price distribution. Furthermore, firms with a star attract more customers (i.e., there is greater market tightness for high quality goods). In equilibrium, buyers are indifferent between buying a good of unknown quality at a low price and high probability or a high quality good at a high price but with a low probability. Finally, we derive the welfare gains from having a rating system. Due to free entry of sellers, sellers do not gain in equilibrium from having a ratings system but buyers may gain or lose from a rating system. The rating system gives them more information about who is a higher quality producer but it is costly – they pay a higher price and face a lower probablity of obtaining a high quality good. But they also gain from paying a lower price if they buy from an unrated seller. We show that this latter effect dominates so a ratings sytem improves welfare..

We then compare the predictions of the model to data obtained on Airbnb pricing in four major cities: Amsterdam, Rome, Miami and San Francisco. We compare the prices and revenues of Superhosts to non-Superhosts, controlling for a range of factors to see if that data supports our model. Controlling for property type, we show that Superhosts charge higher average prices, have a higher occupancy rate and more revenue in each of the four cities we examine.

The paper is organized as follows. Section 2 contains a brief literature review. Section 3 describes the search environment and the seller's pricing decisions, and the welfare benefits of a rating system. Section 4 describes the data we use and presents the empirical findings. Section 5 concludes.

2 Literature Review

The literature on experience goods began with Nelson's (1970) description of the good and how it differs from a pure search good. The main difference is that goods in the search literature have observable attributes once the buyer finds a seller. With an experience good, the buyer does not know its true quality (e..g., taste, comfort, sound or product failure rate) until the good has been bought and consumed. A substantial literature arose that focused on various issues on trade in experience goods such as pricing dynamics, reputation, signalling, advertising and learning about quality. We will address a few key papers in this literature and how it relates to our work.

A common assumption in the literature is that the firm knows the true quality of the good but the consumer does not. Based on the asymmetric information assumpion, Wolinski (1983) studies a firm's decision for choosing the quality level of its good and the role that the price plays in signalling the quality of the good. For a given price, the consumer expects a certain quality level. The firm faces a choice between lowering the quality level (and costs) at the price versus the potential loss of future customers who learn the price does not match the quality. If the potential loss of sales outweighs the cost saving, then in equilibrium, prices reveal the correct quality levels.

A common finding in this literature is that firms have an incentive to lower the price of their good at some point to induce consumers to acquire experience about the good and quality discovery. Shapiro (1983) examines dynamic pricing of experience goods. In his model, agents have an initial perception of the true quality of a good but learn about its true quality by consuming it over time. The case most related to our paper, is when consumers are initially "pessimistic" about the quality of the good. In response, a monopolist producer charges a low price to induce agents to buy the good and then raises the price over time as consumers learn the true quality.[1]

The repeated nature of trades in this type of market also creates heterogeneity in consumer knowledge about the quality of the product. Consequently, a monopolist producer of a good faces two types of consumer markets – a market with experienced consumers and a market with new consumers. Bergemann and Välimäki (2006) study this dynamic pricing problem. In short, a monopolist's current and future pricing decision affects sales to experienced customers and the value of acquiring information by new consumers. By charging the static monoply price, the firm maximizes current profits from experienced customers but deters new customers from buying, thereby reducing its future customer base. Thus, the monopolist faces a tradeoff – by lowering his current price, he loses profits from experienced customers but lowers the cost of information acquisition for new consumers.

There is also a substantial literature on price distributions in the search literature based on the seminal work by Burdett and Judd (1983).[2] The basic idea of these models is that a firm can choose any price but faces a tradeoff – post a higher price and sell fewer units or post a lower price and sell more. In equilibrium, the firm is indifferent between these two strategies and the equilibrium price distribution is non-degenerate.

How does our paper fit into this literature? First, we differ from the existing literature in a key way – there is no asymmetric information about the quality of the good. Producers

[1] A related idea comes from advertising – a firm is willing to incur a cost (lower profits) from advertising if it induces consumers to experience the good and buy it repeatedly. After that prices can rise once the quality of the good is revealed. See Nelson (1974) and Milgrom and Roberts (1986).

[2] See for example Curtis and Wright (2004), Head and Kumar (2005) and Head et al (2012)

in our model are just as uninformed about the quality of the good as the buyer. This eliminates any role for signaling by the producer of the good. However, there is private information about the quality that is experienced by the buyer and the seller wants this information communicated, which occurs via the rating the buyer gives the seller. Second, our model is similar to these other papers in that the producer has an incentive to lower its initial price to increase the probability of a sale. If the match reveals that the producer is of high quality, he then can raise his price and sell more. One key difference is that we can study the welfare gains of a rating system. To our knowledge, this is major difference with the existing literature. Our paper differs from the search literature on price dispersion in that we have imperfect information about the quality of the good. In our model, once quality is revealed the seller can charge a higher price and have higher sales. Finally, most of the papers in the economic literature are purely theoretical whereas we empirically test the predictions of our model using AirBNB pricing data.

Finally, we want to recognize that there is a literature in the marketing and hospitality fields regarding the role of ratings, consumer demand and particularly, pricing of Airbnb rentals.[3] This literature examines the effects of ratings on average price, forecasting of sales or whether they have any impact at all. While most of this literature is purely empirical in nature, Jiang and Wang (2008) and Sun (2012) are exceptions. Both papers study static monopoly pricing in a world in which ratings convey information about quality. Our paper differs from these papers in that is based on a theoretical search model in which the demand a firm faces is endogenous, based on market tightness, and we study dynamic, rather than static, pricing decisions of the seller.

3 A search model with heterogenous quality

3.1 Environment

Time is discrete and goes on forever. The economy is populated by two agent types: buyers and sellers. Sellers produce an indivisible good at no cost, which can be of high or of low quality. Sellers post prices and buyers, who want to consume exactly one unit of the good, attempt to locate sellers. Matching occurs according to a matching function specified below.

Buyers who consume the high-quality good get $u = \varepsilon$, where ε is drawn from a uniform distribution with support $[0, 1]$ and is match-specific. Buyers who consume the low-quality good get $u = 0$. In a match between a buyer and a seller, the buyer first observes the value of ε specific to the match, then the buyer decides whether he wants to acquire the good at

[3] Chen et al (2004) is an early example of the empirical analysis of 'star' rating systems on average price and sale of Amazon products. Koomans (2018) conducts an empirical analysis of how ratings affect pricing attributes of Airbnb rentals. However, there is no theoretical model in her paper.

the posted price. If he agrees to the posted price, the seller produces the good and they separate. After consumption, the buyer exits the market. We assume transferable utility.

Sellers live for two periods and can sell one unit in each period. The probability that a seller is a high-quality seller is x. In the first period of life, the seller's type is unknown to both the seller and the buyer. After production, however, the type of the seller is revealed and the buyer rates the seller. The rating is $R \in \{H, L\}$. We assume that the rating is truthful.

In our environment, sellers can be in three states: unrated in their first period of life, unrated in their second period, and rated in their second period. Without loss in generality, we assume that sellers who receive a rating $R = L$ exit the market, since the rating is public information and no buyer wants to consume a low-quality good. In accordance with the three possible states, three prices are posted. The price p_0 is the price posted by a (unrated) seller in his first period of live. The price p_{00} is the price posted by an unrated seller in his second period of live, and p_1 denotes the price posted by a (high-quality) rated seller in his second period of life.

Buyers observe the three posted prices and direct their search towards one of the three prices so that for each price there is an associated market. In each market, buyers and sellers (who post that particular price) are matched according to a matching function $\mathcal{M}(b, s)$, where b is the measure of buyers and s is the measure of sellers in a particular market. We assume that the matching function has constant returns to scale, and is continuous and increasing with respect to each of its arguments. The measure of buyers is normalized to one. In contrast, the measure of sellers is determined by a free entry condition discussed below.

Let α_0 be the probability of a match for an unrated seller in the first period of his live, α_{00} be the probability of a match for an unrated seller in the second period of his live and α_1 be the probability of a match for a rated seller. Accordingly, we have $\alpha_i = M(s_i, b_i)/s_i$, for $i \in \{0, 1, 00\}$. Let θ_i denote tightness in market i, θ_i is

$$\theta_i = b_i/s_i. \tag{1}$$

The probability of a match for a seller in market i is

$$\alpha_i = m(\theta_i). \tag{2}$$

The probability of a match for a buyer in market i is $\eta_i = M(s_i, b_i)/b_i = M(s_i, b_i)/(\theta_i s_i)$. Thus,

$$\eta_i = m(\theta_i)/\theta_i. \tag{3}$$

As usual, tightness affects positively the probability of a match for a seller, and affects negatively the probability of a match for a buyer. In particular, we assume that $\alpha_i(\theta_i)$ is a strictly increasing and concave function such that $\alpha_i(0) = 0$, $\alpha_i(\infty) = 1$, $\alpha_i'(0) > 0$, $\eta_i(\theta_i) = \alpha_i(\theta_i)/\theta_i$ is strictly decreasing, and $\eta_i(0) = 1$.

3.2 Agents' decisions

In this section, we study the decisions taken by buyers and sellers. Note that there is no private information in a match: Agents are either symmetrically uninformed about the seller's type (in an unrated match) or they are symmetrically informed (in a rated match).

3.2.1 Buyers' acceptance decisions

Buyers get utility $u = \varepsilon$ from consuming the high-quality good and utility $u = 0$ from consuming the low-quality good. Since x is the probability that an unrated seller is a high-quality seller, a buyer who is matched to an unrated seller accepts a posted price if and only if

$$x\varepsilon \geq p_0 \text{ or } x\varepsilon \geq p_{00}. \tag{4}$$

A buyer who is matched to a rated seller accepts a posted price if and only if

$$\varepsilon \geq p_1. \tag{5}$$

Accordingly, the expected utilities of searching in the two unrated markets are

$$\eta_0 \int_{p_0/x}^{1} (x\varepsilon - p_0)\, d\varepsilon \text{ and } \eta_{00} \int_{p_{00}/x}^{1} (x\varepsilon - p_{00})\, d\varepsilon. \tag{6}$$

On the left in (6), with probability η_0 the buyer has a match with an unrated young seller and the match-specific utility ε is learnt. For all $\varepsilon \geq p_0/x$ the buyer accepts the posted price p_0 and the good is produced. The buyer then gets utility ε with the probability x that the good is high quality minus the price p_0. The term on the right in (6) has a similar interpretation.

In the rated market the expected utility is

$$\eta_1 \int_{p_1}^{1} (\varepsilon - p_1)\, d\varepsilon. \tag{7}$$

In (7), with probability η_1 the buyer has a match with a rated seller and learns the match-specific utility ε. For all $\varepsilon \geq p_1$ the buyer accepts the posted price p_1 and the good is produced. The buyer then gets utility ε with certainty since the seller is high quality, minus the price p_1.

By taking into account the expected utilities stated in (6) and (7), buyers direct their search towards the market that yields the highest expected utility.

3.2.2 Sellers' price posting decisions

In the unrated market, buyers accept a trade if (4) holds, and in the rated market, they accept a trade if (5) holds. Accordingly, the sellers' value functions in the three markets are, respectively,

$$V_0 = \alpha_0 \int_{p_0/x}^{1} d\varepsilon \left(p_0 + xV_1\right) + \left(1 - \alpha_0 \int_{p_0/x}^{1} d\varepsilon\right) V_{00}, \tag{8}$$

$$V_1 = \alpha_1 \int_{p_1}^{1} d\varepsilon p_1 = \alpha_1 \left(1 - p_1\right) p_1, \tag{9}$$

and

$$V_{00} = \alpha_{00} \int_{p_{00}/x}^{1} d\varepsilon p_{00} = \alpha_{00} x \left(1 - p_{00}/x\right) p_{00}/x. \tag{10}$$

In (8) the expected value of being an unrated seller in the first period of life is as follows. With probability α_0 this seller is matched with a buyer and the value of ε becomes known. Then the seller is paid the price p_0 if the buyer gets a non-negative expected payoff; i.e., for all $\varepsilon \in [p_0/x, 1]$. If the good produced is high quality (with probability x) the seller receives a rating $R = H$ and enters the rated market in the following period, with associated expected payoff V_1. If the good produced is low quality (with probability $1-x$) the seller receives a rating $R = L$ and exits the economy. If the seller in the first period of life has a match that yields a low value of ε or does not have a match, the seller remains unrated and enters the unrated market for sellers in their second period of life, with associated expected payoff V_{00}. Equations (9) and (10) have similar interpretations, except that both rated sellers and unrated old sellers exit the economy with certainty after participating in the respective markets, since they are all in their second period of life.

In each state $i = \{0, 00, 1\}$ the sellers choose price p_i in order to maximize their life-time utility; i.e. they chose p_i such that the right hand-sides of (8)-(10) are maximized.

The first-order conditions on p_0, p_1, and p_{00} are, respectively,

$$p_0 = (x/2)\left[1 - (V_1 - V_{00}/x)\right], \tag{11}$$

$$p_1 = 1/2, \tag{12}$$

and

$$p_{00} = x/2. \tag{13}$$

Replacing p_1 in (9) and p_{00} in (10), we obtain the following expressions for the value functions

$$V_1 = \alpha_1/4 \text{ and } V_{00} = x\alpha_{00}/4. \tag{14}$$

Using (14), we can rewrite (11) as follows

$$p_0 = \frac{x\left[1 - (\alpha_1 - \alpha_{00})/4\right]}{2}. \tag{15}$$

If the arrival rates for old rated and unrated sellers are equal; i.e., $\alpha_1 = \alpha_{00}$, then $p_0 = p_{00}$ and all unrated sellers set the same price. However, if old rated sellers face higher arrival rates than unrated sellers; i.e., $\alpha_1 > \alpha_{00}$, then $p_0 < p_{00} = x/2$, and the price set by unrated sellers is lower.

3.3 Rating Equilibrium

3.3.1 Free entry condition

As usual, the free entry condition is $V_0 = k$, where k is a fixed utility cost of entering the market. The value function V_0 simplifies as follows

$$V_0 = \alpha_0 x\left(1 - p_0/x\right)^2 + x\alpha_{00}/4. \tag{16}$$

Then free entry implies

$$k = \alpha_0 x\left(1 - p_0/x\right)^2 + x\alpha_{00}/4.$$

Using (11) to replace p_0 we get

$$k = \alpha_0 x\left[1 + (\alpha_1 - \alpha_{00})/4\right]^2/4 + x\alpha_{00}/4, \tag{17}$$

an equilibrium equation in the arrival rates for sellers.

3.3.2 Buyers' directed search

Buyers observe prices set as in (12), (13), and (15). Moreover, they correctly anticipate the queue length in each market and direct their search to the market which promises the highest expected utility. In equilibrium, they must be indifferent between the three options. From (6) and (7), buyers are indifferent if

$$\eta_0 \int_{p_0/x}^{1} (x\varepsilon - p_0)\, d\varepsilon = \eta_{00} \int_{p_{00/x}}^{1} (x\varepsilon - p_{00})\, d\varepsilon = \eta_1 \int_{p_1}^{1} (\varepsilon - p_1)\, d\varepsilon. \tag{18}$$

Using (12) and (13), the second equality of (18) simplifies to

$$\eta_1 = \eta_{00} x. \tag{19}$$

Thus, if $x < 1$, for a buyer the probability of a match is smaller in the rated market than in the unrated one. From (18), buyers's expected utility is $\eta_1/8 = \eta_{00}x/8$ since $p_{00} = x/2$ and $p_1 = 1/2$.

The first equality of (18) can be written as follows

$$\eta_0 \left(1 - p_0/x\right)^2 = \eta_{00} \left(1 - p_{00}/x\right)^2.$$

Using (11) and (13), this expression can be simplified as

$$\eta_0 \left[1 + \left(\alpha_1 - \alpha_{00}\right)/4\right]^2 = \eta_{00}. \tag{20}$$

The following Lemma characterises prices and matching probabilities across markets.

Lemma 1 *In any equilibrium with $x < 1$, prices satisfy $p_1 > p_{00} > p_0$ and matching probabilities satisfy $\alpha_1 > \alpha_0 > \alpha_{00}$ (or $\eta_1 < \eta_0 < \eta_{00}$). For $x = 1$, we have $p_1 = p_{00} = p_0$ and $\alpha_1 = \alpha_0 = \alpha_{00}$ (or $\eta_1 = \eta_0 = \eta_{00}$).*

Proof. From (19), for a buyer, the probability of a match is smaller in the rated market than in the unrated one if $x < 1$. Accordingly, $\theta_1 > \theta_{00}$; i.e., the ratio of buyers to sellers is larger in market 1 than in market 00. This implies that $\alpha_1 > \alpha_{00}$ if $x < 1$. Then, (15) yields $p_0 < p_{00} = x/2 < p_1 = 1/2$. From (20), since $\alpha_1 > \alpha_{00}$ it follows that $\eta_0 < \eta_{00}$, which implies that $\theta_0 > \theta_{00}$ and therefore $\alpha_0 > \alpha_{00}$. If $x = 1$, then from (19) we have $\eta_1 = \eta_{00}$ and so $\alpha_1 = \alpha_{00}$ implying that $p_0 = p_{00} = p_1 = 1/2$ and $\eta_0 = \eta_{00}$. ■

Lemma 1 shows that in any equilibrium the arrival rate for rated sellers is higher than the arrival rate for unrated ones, and that the arrival rate for unrated sellers in their first period of life is higher than the arrival rate for unrated sellers in their second period of life. Furthermore, we find that rated sellers post a higher price than unrated ones, and unrated sellers in their first period of life post the lowest price. In particular, they post a lower price than the unrated sellers in their second period of life. This shows that sellers attempt to get a rating even when they are unaware of their type.

With the above equations, we are able to compute the equilibrium as stated in the following definition.

Definition 1 *A rating equilibrium are prices* p_i *and tightnesses* θ_i*, for* $i \in \{0, 1, 00\}$ *that solve (12), (13), (15), (17), (19), and (20).*

Proposition 1 *There is* $\hat{x} < 1$ *such that if* $x \geq \hat{x}$*, then a rating equilibrium with* $\theta_i > 0$ *and* $p_i > 0$ *for all* i *exists and is unique. In this equilibrium welfare is decreasing in* k*.*

Proof. See Appendix. ■

From Proposition 1, if the average quality of sellers x is sufficiently high there is a unique solution for tightnesses θ_1, θ_0 and θ_{00} that define the rating equilibrium. To understand why, notice first from (19) that for a given x tightnesses θ_1 and θ_{00} are positively correlated, meaning that the sellers' matching probabilities α_1 and α_{00} are also positively correlated. The free-entry condition for sellers (17) then implies that, for a given entry cost k, if tightness θ_{00} decreases, then α_{00} and α_1 decrease, and hence for sellers to enter the market it must be that θ_0 increases so that α_0 increases as well. Therefore, the sellers' free-entry condition (17) defines a negative relationship between θ_0 and θ_{00}. From (20), for buyers to be indifferent between the market of unrated sellers in their first period of life and the market of unrated sellers in their second period of life, if θ_{00} increases and hence η_{00} decreases it must be that $\eta_0 \left[1 + (\alpha_1 - \alpha_{00})/4\right]^2$ decreases as well. However, a decrease in η_{00} may be offset by a decrease in η_0 or by a decrease in $(\alpha_1 - \alpha_{00})$, depending on the elasticity of the function α_i. If x is sufficiently high, (19) implies that α_1 and α_{00}, as well as the respective elasticities, are sufficiently close so that the overall elasticity of $(\alpha_1 - \alpha_{00})$ is relatively small. In that case, (20) unambiguously defines a positive relationship between η_{00} and η_0, and hence a positive relationship between θ_{00} and θ_0. Therefore, equations (17) and (20), together with (19), yield a unique solution for tightnesses θ_0, θ_{00}, and θ_1 which in turn from (15) determine p_0.

Proposition 1 also states that for $x \geq \hat{x}$ an increase in the entry cost k is welfare worsening, because buyers' expected utility of searching decreases with k (sellers' expected utility is always zero in equilibrium). Intuitively, a greater entry cost deters sellers' entry and thereby for a given price reduces the buyers' matching probability.

3.3.3 Do sellers want to be rated?

So far we have assumed that it is public knowledge if a seller has sold a unit. Suppose now that a seller can opt out from being rated but he still posts the same prices as all other sellers. Would he opt out? In the first period of his life he gets the same matching probability and charges the same price as all other agents. However his continuation payoff is different since by hiding this information he can go to the unrated market even if he is a low-quality seller. Accordingly, his life-time utility denoted by $\bar{V}_0$ is

$$\bar{V}_0 = \alpha_0 \int_{p_0/x}^{1} p_0 d\varepsilon + V_{00} \tag{21}$$

Using (14), (21) becomes

$$\bar{V}_0 = \alpha_0 \left(1 - p_0/x\right) p_0 + x\alpha_{00}/4 \tag{22}$$

Comparing the equilibrium expected payoff in (16) with $\bar{V}_0$, it follows that the seller has no incentive to opt out if $(1 - p_0/x) \geq p_0/x$ which simplifies to $p_0 \leq x/2$. From (15), $p_0 \leq x/2$ holds if $\alpha_{00} \leq \alpha_1$, which is the case in equilibrium. Therefore, the rating system is incentive-compatible: sellers prefer taking the chance of getting a rating that reveals their quality even at the risk of exiting the market before their second period of life. Notice that incentive compatibility holds irrespective of the value of x. The reason is that, even if x is relatively low and hence the chance for sellers of being of high quality is small, a low x also means that the price that sellers can charge in the unrated market is also low compared to the price that they can charge in the rated market. Therefore sellers are better off by seeking to get a rating for all values of x.

3.4 Equilibrium in the absence of ratings

Here we calculate the equilibrium in the absence of ratings to see whether ratings improve the allocation. The free entry condition implies that for sellers nothing changes. However, buyers can be better or worse off in the absence of ratings. The only equilibrium for old sellers is pooling since there is no cost of producing the goods and so there can be no separating equilibrium.

In the absence of ratings, there are only two states, and hence two value functions $\tilde{V}_0$ and $\tilde{V}_1$ for old and young sellers, respectively, where the tilde is used to indicate the variables in the economy without ratings. These value functions are

$$\tilde{V}_0 = \tilde{\alpha}_0 \int\limits_{\tilde{p}_0/x}^{1} d\varepsilon \tilde{p}_0 + \tilde{V}_1$$

and,

$$\tilde{V}_1 = \tilde{\alpha}_1 \int\limits_{\tilde{p}_1/x}^{1} d\varepsilon \tilde{p}_1 = \tilde{\alpha}_1 x \left(1 - \tilde{p}_1/x\right) \tilde{p}_1/x.$$

By combining the two equations above we can rewrite $\tilde{V}_0$ as follows

$$\tilde{V}_0 = \tilde{\alpha}_0 x \left(1 - \tilde{p}_0/x\right) \tilde{p}_0/x + \tilde{\alpha}_1 x \left(1 - \tilde{p}_1/x\right) \tilde{p}_1/x$$

The first-order conditions for $\tilde{p}_0$ and $\tilde{p}_1$ that follow from maximizing $\tilde{V}_0$ and $\tilde{V}_1$ are

$$\tilde{p}_0 = \tilde{p}_1 = x/2$$

Therefore, the free entry condition implies that

$$k = \tilde{V}_0 = \tilde{\alpha}_0 x/4 + \tilde{\alpha}_1 x/4. \tag{23}$$

Buyers must be indifferent between the two markets, hence the following condition must hold

$$\tilde{\eta}_0 \int\limits_{\tilde{p}_0/x}^{1} \left(x\varepsilon - \tilde{p}_0\right) d\varepsilon = \tilde{\eta}_1 \int\limits_{\tilde{p}_1/x}^{1} \left(x\varepsilon - \tilde{p}_1\right) d\varepsilon$$

Since the prices $\tilde{p}_0$ and $\tilde{p}_1$ are equal, the probabilities $\tilde{\eta}_0$ and $\tilde{\eta}_1$ are also equal, and therefore tightness in the market for young sellers $\tilde{\theta}_0$ equals tightness in the market for old sellers $\tilde{\theta}_1$. Using (23), this implies that

$$\tilde{\alpha}_0 = \tilde{\alpha}_1 = 2k/x.$$

In order to compare welfare in the economy without ratings with welfare in the economy with ratings, we rewrite equation (17) to obtain

$$\frac{4k}{x} - \alpha_0 \left[1 + \left(\alpha_1 - \alpha_{00}\right)/4\right]^2 - \alpha_{00} = 0$$

where $\alpha_1 > \alpha_{00}$ and $\alpha_0 > \alpha_{00}$. Therefore $\alpha_{00} < (4k/x)/2 = 2k/x$. Comparing α_{00} with $\bar{\alpha}_0$, we have that $\alpha_{00} < \bar{\alpha}_0 = 2k/x$, which implies that $\theta_{00} < \bar{\theta}_0$ and therefore $\eta_{00} > \bar{\eta}_0$. In addition, recall that $p_{00} = x/2$, so that $p_{00} = \bar{p}_0 = x/2$. Since the matching probability for buyers is higher in market 00 in the economy with ratings than in market 0 in the economy without ratings, while the price and the probability of acquiring a high-quality good are identical in both cases, buyers' utility is higher in the former case. Since buyers' utility is the same across markets within the same economy, it follows that buyers are better off in the economy with ratings. The following proposition summarises this result.

Proposition 2 *In an economy with ratings welfare is higher than in an economy without ratings.*

4 Price setting and market tightness on Airbnb

We illustrate the model presented in Section 3 with data from Airbnb, an online platform for rentals that allows hosts and guests to be matched. Hosts can post their listings, and guests can search for rentals that best suit their preferences. As part of its intermediation services, Airbnb encourages guests to rate their trip experiences on a 1-5 scale and leave reviews. The average star rating and the reviews are then made visible online for each listing.

We use data on 4 major cities: Amsterdam, Rome, Miami and San Francisco. Inspecting these cities is interesting because they are all highly touristic and, at the same time, they allow us to see whether the predicted patterns hold in sufficiently diverse locations. The key aspect of the data that we explore is the role of the superhost status that hosts may acquire through Airbnb. The superhost status can be interpreted as a label that high-quality sellers acquire based on their selling experience. There are four requirements that hosts must meet to become a superhost on Airbnb. First, superhosts must not cancel reservations, unless there are extenuating circumstances. Second, superhosts must maintain a response rate of at least 90% when they are contacted by guests. Third, superhosts must receive 5 stars in at least 80% of their reviews, and they must receive reviews by no less than half of their guests.[4] Finally, superhosts must host at least 10 trips in the past year. In exchange for meeting the above criteria, an important benefit of becoming a superhost is the "Superhost badge" delivered by Airbnb. This badge appears on the superhost' profile and listing pages and precisely certifies that the host has complied with the superhost' requirements.

The dataset contains key variables with one value for each rental listed on Airbnb. These variables are computed on a yearly basis: average daily rate over the last twelve months, number of bookings in the last twelve months, occupancy rate over the last twelve months, and annual revenue. The dataset also contains information about the physical

[4]The minimum rating requirement has been updated by Airbnb recently, but it was computed as described at the moment the data was collected.

characteristics of the rentals such as the number of bedrooms, the number of bathrooms, the maximum number of guests, and the neighborhood. The variables that are informative about the quality of the rental (the cumulative overall rating, the superhost status, and the number of reviews) are unfortunately only available for the last date at which the rental was located on the Airbnb website. Therefore this data allows us to make a cross-sectional comparison among the rentals.[5]

We restrict the sample as follows. We keep only listings that are entire apartments, and drop houses, villas, bungalows, bed & breakfasts, etc., as well as shared or private rooms.[6] We drop a small number of rentals that appear to have 0 bathrooms. We drop observations for which the average daily rate is missing or is higher than 3 standard deviations, or those with missing superhost status, neighborhood, or city in the case of the metropolitan areas of Miami and San Francisco.[7] Finally, we restrict attention to rentals with at least 10 bookings which are most likely run as a professional activity or business, compared to rentals that only host a few trips per year. This allows us to exclude rentals which target only dates in high season or specific events in the city since their rates could be disproportionately high. This leaves us with 4709 observations for Amsterdam, 4183 for Rome, 2073 for San Francisco, and 4272 for Miami.

Table 1: Descriptive statistics. Rentals' characteristics ($\geq$ 10 bkngs)

	Amsterdam	Rome	San Francisco	Miami
1 bedroom	57.9%	46.6%	47.1%	45.9%
2 bedrooms	28.1%	33.4%	25.6%	26.8%
1 bathroom	86.6%	71.4%	82.6%	70.1%
2 guests max.	45.9%	14.5%	36.2%	16.0%
3 guests max.	11.8%	11.3%	12.3%	12.7%
4 guests max.	35.3%	36.5%	28.3%	37.5%

Source: Authors' calculations from AIRDNA data

Table 1 presents descriptive statistics on the rentals' physical characteristics for the

[5]For most rentals, the last date at which information was collected is July 2016 (once we restrict the dataset as described below, these rentals represent 63.6% in Amsterdam, 80.9% in Rome; 63% in San F., and 64% in Miami). For the remaining rentals, the last information is older than July 2016, presumably because these rentals were no longer active as of this date (although the last information on almost all rentals was collected between January 1, 2016, and mid-July, 2016; the percentages are 94.7% for Amsterdam, 98% for Rome, 95.4% for San F., and 97.5% for Miami). Notice that we need to keep these rentals in our sample in order to take into account the prices set by hosts whose rentals are presumably low quality and exit the market.

[6]In terms of listing type, entire homes or apartments represent 80.5% of observations in Amsterdam, 67.1% in Rome, 54.3% in San Francisco, and 70.3% in Miami. In terms of property type, apartments represent 81.3% of observations in Amsterdam, 73.9% in Rome, 52.2% in San Francisco, and 65.5% in Miami.

[7]For Miami there is no information about neighborhood.

Table 2: Descriptive statistics ($\geq$ 10 bkngs)

	mean	sd	min	max
Amsterdam (4709 obs.)				
average daily rate	163.1	61.7	44.6	415.0
occupancy rate	0.65	0.20	0.09	1
nr bookings	26.3	19.1	10	138
annual revenue	15871	11606	1653	140894
Rome (4183 obs.)				
average daily rate	140.5	60.9	34.4	447.0
occupancy rate	0.57	0.19	0.08	1
nr bookings	34.6	22.7	10	147
annual revenue	17170	11147	565	88811
San Francisco (2073 obs.)				
average daily rate	233.3	117.7	39.2	980.0
occupancy rate	0.69	0.18	0.09	1
nr bookings	31.9	22.4	10	143
annual revenue	29788	22807	1242	208296
Miami (4272 obs.)				
average daily rate	191.4	103.0	42.1	842.9
occupancy rate	0.59	0.18	0.10	1
nr bookings	26.8	18.3	10	134
annual revenue	21146	13506	1315	142576

Source: Authors' calculations from AIRDNA data

resulting sample. In the four cities most rentals have 1 bedroom and 1 bathroom, and can host a maximum of 2 guests (Amsterdam and San Francisco) or 4 guests (Rome and Miami). Table 2 presents descriptive statistics on the average daily rate, the occupancy rate, the number of bookings and the revenue for the past year (the average daily rate and the annual revenue are computed in U.S. dollars for the four cities.)

Table 3 presents the fraction of superhosts in our sample that goes from 10.3% in Miami to 17.8% in San Francisco, and is thus similar across the four cities.[8]

The model presented in Section 3 predicts that old or experienced high-quality sellers, that we see as being the superhosts in the Airbnb sample, charge higher prices, have a higher probability of sale, and earn a higher revenue. In order to see whether these predictions hold, we compare superhosts and non-superhosts based on their average daily rates, their occupancy rates, and their annual revenue. First, we make a comparison between superhosts and non-superhosts without controlling for rentals' characteristics.

[8]Considering the same sample before excluding rentals with less than 10 bookings in the past year, we have 8.4% of superhosts in Amsterdam (out of 10604 obs.), 9.5% in Rome (out of 6784 obs.), 10.4% (out of 4295 obs.) in San Francisco and 6.0% (out of 10329 obs.) in Miami.

Table 3: Fraction of superhosts ($\geq$ 10 bkngs)

	superhosts (%)	non-superhosts (%)
Amsterdam (4709 obs.)	15.1	84.9
Rome (4183 obs.)	14.0	86.0
San Francisco (2073 obs.)	17.8	82.2
Miami (4272 obs.)	10.3	89.8

Source: Authors' calculations from AIRDNA data.

Table 4: Superhosts and non-superhosts ($\geq$ 10 bkngs)

	Superhost (mean)	Non-superhost (mean)	Diff.	Superhost (mean)	Non-superhost (mean)	Diff.
		Amsterdam			San Francisco	
avge. daily rate	168.2	162.2	6.0*	239.5	232.0	7.5
occupancy rate	0.69	0.64	0.05***	0.74	0.68	0.06***
nr bookings	30.2	25.6	4.69***	39.3	30.4	9.0***
ann. revenue	19061	15304	3757***	37521	28113	9407***
obs.	711	3998	4709	369	1704	2073
		Rome			Miami	
avge. daily rate	141.4	140.4	1.0	177.8	192.9	-15.1**
occupancy rate	0.66	0.56	0.11***	0.66	0.58	0.08***
nr bookings	43.6	33.1	10.55***	32.5	26.1	6.34***
ann. revenue	22937	16235	6702***	24804	20728	4077***
obs.	584	3599	4183	438	3834	4272

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Source: Authors' calculations from AIRDNA data

Table 4 presents the average daily rate, the occupancy rate, and the annual revenue for all rentals in our restricted sample, sorted by superhost status. The number of bookings in the last twelve months is also presented. In all cities, superhosts earn on average a higher annual revenue than non-superhosts. The reason is that the average number of bookings per year is considerably higher for superhosts, while their average daily rate is higher or sufficiently similar.

The overall rating (average number of stars) and the average number of reviews received by superhosts and non-superhosts are presented in Table 5 (for the same subset of rentals as in previous tables). Consistently with the superhosts' requirements, the average values of these two variables are systematically higher for superhosts. Table 5 also presents hosts' average response time and average response rate. As expected, superhosts are more prone to respond and respond faster to their guests than non-superhosts. In addition, superhosts include more photos in their listings than non-superhosts.

In order to control for specific characteristics of the rentals, we run standard OLS

Table 5: Superhosts and non-superhosts ($\geq$ 10 bkngs)

	Superhost (mean)	Non-superhost (mean)	Diff.	Superhost (mean)	Non-superhost (mean)	Diff.
		Amsterdam			San Francisco	
overall rating	4.85	4.59	0.27***	4.86	4.63	0.23***
nr reviews	39.0	24.5	14.5***	55.0	34.3	20.7***
resp. rate	97.5	92.7	4.8***	98.1	91.9	6.2***
resp. time (min)	186.5	269.3	-82.8***	134.6	242.7	-108.1***
nr photos	20.6	17.8	2.8***	20.2	17.2	3.0***
obs.	711	3998	4709	369	1704	2073
		Rome			Miami	
overall rating	4.8	4.5	0.3***	4.83	4.47	0.36***
nr reviews	55.4	34.1	21.3***	36.7	21.8	14.9***
resp. rate	99.1	96.6	2.5***	98.6	92.5	6.0***
resp. time (min)	65.9	124.5	-58.6***	62.3	196.8	-134.5***
nr photos	29.1	25.5	3.6***	25.8	22.3	3.4***
obs.	584	3599	4183	438	3834	4272

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Source: Authors' calculations from AIRDNA data

regressions for each of the four cities. We set as dependent variables the average daily rate (in log), the occupancy rate in the last twelve months and the annual revenue (in log). Notice that annual revenue is not mechanically implied by the occupancy rate and the average daily rate because it also depends on the number of bookings and the number of reservation days per booking. Our aim is to test whether the coefficient associated with superhost status is statistically significant. In the regressions we include as controls the following individual characteristics: size of the apartment (number of bedrooms, number of bathrooms, maximum number of guests), overall rating, number of photos in the listing, number of reviews, and dummies for the neighborhood and city where rentals are located. Since number of bedrooms, number of bathrooms, and maximum number of guests are highly correlated, we drop the number of bedrooms and/or the number of bathrooms. Airbnb' hosts can enable the "instant book" option that allows guests to directly book a rental through the platform (which checks for availability for the required dates) without prior contact with the host. Since this option may allow hosts to post higher prices and get more bookings similarly as the superhost status, we include it in the regression.

With these regressions, we confirm that the superhost status is closely linked to hosts' pricing decisions and their probability of receiving trips. Tables 6-8 present results for Amsterdam (the tables for the other cities are presented in the appendix). The coefficient for superhost status is statistically significant in the regression with the average daily rate earned by hosts as the dependent variable. The size of the apartment (maximum

Table 6: Amsterdam. Dependent variable: average daily rate (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.185***	0.185***	0.179***	0.164***	0.160***	0.159***
	(0.0145)	(0.0144)	(0.0142)	(0.0142)	(0.0139)	(0.0139)
Max Guests	0.133***	0.135***	0.140***	0.135***	0.136***	0.136***
	(0.00325)	(0.00325)	(0.00325)	(0.00327)	(0.00321)	(0.00321)
superhost		0.0706***	0.0430***	0.0341**	0.0522***	0.0521***
		(0.0107)	(0.0109)	(0.0108)	(0.0107)	(0.0107)
Overall Rating			0.112***	0.101***	0.104***	0.107***
			(0.0117)	(0.0116)	(0.0114)	(0.0115)
Number of Photos				0.00377***	0.00448***	0.00447***
				(0.000393)	(0.000389)	(0.000389)
Number of Reviews					-0.00150***	-0.00151***
					(0.000110)	(0.000111)
instantbook						0.0148
						(0.0112)
Constant	4.338***	4.332***	3.616***	3.651***	3.684***	3.669***
	(0.118)	(0.118)	(0.158)	(0.157)	(0.154)	(0.154)
Observations	4701	4701	4612	4612	4612	4612
Adjusted R^2	0.421	0.426	0.442	0.453	0.474	0.474

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

number of guests and number of bathrooms), the overall rating, and the number of photos included in the listing also have statistically significant positive coefficients (the coefficient for the instant book option is not statistically significant). The coefficient on the number of reviews is negative and statistically significant although it is economically small. A drawback with the number of reviews is that it is highly correlated with the number of bookings (the unconditional correlation is more than 0.7 in the four cities). Therefore the negative coefficient of this variable is capturing the negative correlation, all other things being equal, between number of bookings and rates. In the other cities, regressions give similar results. However, in Rome and Miami, once we include overall rating as predictor the coefficient for superhost status becomes non-significant in some specifications. This is natural, since the roles of overall rating and the superhost status overlap to a large extent.

Table 7 displays the regression with the occupancy rate in the last twelve months as the dependent variable for Amsterdam. While controlling for the size of the apartment,

Table 7: Amsterdam. Dependent variable: occupancy rate ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.0489***	0.0501***	0.0519***	0.0483***	0.0474***	0.0458***
	(0.0109)	(0.0108)	(0.0108)	(0.0108)	(0.0107)	(0.0107)
Max Guests	0.0138***	0.0160***	0.0210***	0.0202***	0.0169***	0.0170***
	(0.00280)	(0.00280)	(0.00288)	(0.00288)	(0.00288)	(0.00287)
logadr	-0.209***	-0.217***	-0.234***	-0.240***	-0.218***	-0.220***
	(0.0108)	(0.0108)	(0.0110)	(0.0111)	(0.0112)	(0.0112)
superhost2		0.0576***	0.0419***	0.0394***	0.0288***	0.0286***
		(0.00789)	(0.00814)	(0.00815)	(0.00815)	(0.00812)
Overall Rating			0.0615***	0.0588***	0.0549***	0.0622***
			(0.00881)	(0.00882)	(0.00875)	(0.00884)
Number of Photos				0.00117***	0.000698*	0.000649*
				(0.000299)	(0.000300)	(0.000299)
Number of Reviews					0.000820***	0.000778***
					(0.0000854)	(0.0000856)
instantbook						0.0432***
						(0.00850)
Constant	1.415***	1.443***	1.038***	1.071***	0.975***	0.936***
	(0.0989)	(0.0985)	(0.125)	(0.125)	(0.124)	(0.124)
Observations	4701	4701	4612	4612	4612	4612
Adjusted R^2	0.091	0.101	0.111	0.114	0.131	0.136

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the price (average daily rate in log), the overall rating and the number of photos, the coefficient on the superhost status is positive and statistically significant. As expected, the coefficients for the maximum number of guests, the overall rating, the number of photos, the number of reviews and the instant book option are all positive, and the one for the price is negative, and all are statistically significant.

Table 8: Amsterdam. Dependent variable: annual revenue (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.210***	0.208***	0.198***	0.154***	0.178***	0.171***
	(0.0330)	(0.0326)	(0.0328)	(0.0325)	(0.0292)	(0.0290)
Max Guests	0.139***	0.145***	0.152***	0.136***	0.132***	0.132***
	(0.00743)	(0.00736)	(0.00751)	(0.00750)	(0.00673)	(0.00669)
superhost		0.256***	0.228***	0.202***	0.109***	0.108***
		(0.0242)	(0.0251)	(0.0248)	(0.0224)	(0.0223)
Overall Rating			0.0905***	0.0578*	0.0413	0.0702**
			(0.0269)	(0.0266)	(0.0239)	(0.0241)
Number of Photos				0.0111***	0.00741***	0.00719***
				(0.000901)	(0.000816)	(0.000812)
Number of Reviews					0.00769***	0.00753***
					(0.000231)	(0.000231)
instantbook						0.172***
						(0.0234)
Constant	8.605***	8.584***	7.570***	7.673***	7.505***	7.331***
	(0.270)	(0.267)	(0.365)	(0.359)	(0.322)	(0.321)
Observations	4701	4701	4612	4612	4612	4612
Adjusted R^2	0.167	0.186	0.192	0.217	0.370	0.377

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

From table 8, being superhost is associated with a higher revenue, with statistically significant coefficients (at the 1% level) in all specifications. Economically, the correlation is large even in the specification that includes size of the rental, overall rating, number of photos, number of reviews and the instant book option. All other things being equal, being superhost is associated with earning a revenue that is typically more than 10% higher.

5 Conclusion

We have presented a search model where ratings allow buyers to get informed about the quality of goods. The ratings affects the sellers' pricing behavior, since they optimally set prices in order to maximize the chances of getting a rating. We have illustrated our model with Airbnb data, and shown that the correlations between prices, revenue and probability across Airbnb rentals are those predicted by our theoretical model. Possible extensions in future research include studying richer rating settings such as double-sided ratings systems, where ratings are posted both by sellers and buyers.

References

- Bergeman, D. and J. Välimäki (2006). “Optimal Pricing of New Experience Goods”, *Journal of Political Economy*, vol. 114 no. 41, 713-743.
- Burdett, K. and K. Judd (1983). “Equilibrium Price Dispersion”, *Econometrica*, 51 (4), 955-969.
- Chen, P., S. Wu and J. Yoon (2004). “The Impact of Online Recommendations and Consumer Feedback on Sales” ICIS 2004 Proceedings, 58, 711-723.
- Curtis, E. and R. Wright (2004). “Price Setting, Price Dispersion and the Value of Money, or the Law of Two Prices” *Journal of Monetary Economics,* 51, 1599-1621.
- Head, A., L. Liu, G. Menizo and R.Wright (2012). “Sticky Prices: A New Monetarist Approach”, *Journal of the European Economic Association,* 10 (5), 939-973.
- Head, A. and A. Kumar (2005). “Price Dispersion, Inflation, and Welfare”, *International Economic Review*, 46 (2), 533-572.
- Jiang, B. and B. Wang (2008). “Impact of Consumer Reviews and Ratings on Sales, Prices, and Profits: Theory and Evidence”, ICIS 2008 Proceedings, 141, 1-17.
- Koomans, Marije (2018) ”The Price Determinants of Airbnb in Rotterdam”, HTH Working Paper Series Vol.2 No.1, May.
- Nelson, P. (1970). “Information and Consumer Behaviour”, *Journal of Political Economy* 87 (2), 311-329.
- Shapiro, C. (1983). “Optimal Pricing of Experience Goods”, *The Bell Journal of Economics,* 14(2), 497-507.
- Wolinsky, A. (1983). “Prices as Signals of Product Quality”, *The Review of Economic Studies*, 50 (4), 647-658.

Appendix A. Proofs

Proof of Proposition 1. In the first part of the proof, we prove the existence and uniqueness of the rating equilibrium for $x \geq \hat{x}$ with $\hat{x} < 1$. In the second part of the proof, we show that in this rating equilibrium welfare decreases with k.

Equilibrium tightnesses solve (17)-(20), which we restate here for ease of reference:

$$k - \alpha_0 x \left[1 + (\alpha_1 - \alpha_{00})/4\right]^2/4 - x\alpha_{00}/4 = 0 \tag{24}$$

$$\eta_1 - \eta_{00} x = 0 \tag{25}$$

$$\eta_0 \left[1 + (\alpha_1 - \alpha_{00})/4\right]^2 - \eta_{00} = 0 \tag{26}$$

Recall that (24) is the free-entry condition by sellers, (25) is the buyers' indifference condition between market 00 and market 1, and (26) is the buyers' indifference condition between market 0 and market 00. From (25), tightnesses θ_1 and θ_{00} move in the same direction. Thus, $\theta_1(\theta_{00})$ is an increasing function of θ_{00}. We can then express everything in terms of θ_0 and θ_{00} as follows

$$4k/x - \alpha_0(\theta_0)\left\{1 + \left[\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})\right]/4\right\}^2 - \alpha_{00}(\theta_{00}) = 0 \tag{27}$$

$$\eta_0(\theta_0)\left\{1 + \left[\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})\right]/4\right\}^2 - \eta_{00}(\theta_{00}) = 0 \tag{28}$$

Equations (27) and (28) define both a relationship between θ_{00} and θ_0. To prove the existence and uniqueness of a solution for θ_{00} and θ_0, we proceed as follows. First, we show that equation in (27) defines θ_0 as a monotonically decreasing function of θ_{00}. Second, we show that equation (28) defines θ_0 as a monotonically increasing function of θ_{00}. Third, we show that the curve defined in equation (27) has a greater intercept value of θ_0 (for $\theta_{00} = 0$) than the curve defined in equation (28). We then conclude that the two curves have a unique intersection for $\theta_0, \theta_{00} > 0$.

Denote the slope of the function $\theta_0(\theta_{00})$ implicit in (27) as $d\theta_0/d\theta_{00}$ and the slope of the function $\theta_0(\theta_{00})$ implicit in (28) as $\hat{d}\theta_0/\hat{d}\theta_{00}$. By using the implicit function theorem, we totally differentiate (27) to obtain

$$\frac{d\theta_0}{d\theta_{00}} = -\frac{-2A\alpha_0\left[\alpha_1'(d\theta_1/d\theta_{00}) - \alpha_{00}'\right]/4 - \alpha_{00}'}{-\alpha_0' A^2}$$

where $A = 1 + (\alpha_1 - \alpha_{00})/4$ and $1 < A < 5/4$ since $\alpha_1 > \alpha_{00}$. Rearranging, this equation can be rewritten as

$$\frac{d\theta_0}{d\theta_{00}} = -\frac{A\alpha_0\alpha_1'(d\theta_1/d\theta_{00})/2 + \alpha_{00}'(1 - A\alpha_0/2)}{\alpha_0' A^2}. \tag{29}$$

In (29), the denominator and the numerator are both positive. It follows that $d\theta_0/d\theta_{00} < 0$ for all θ_{00} in (29); hence (27) defines θ_0 as a monotonically decreasing function in θ_{00}.

Next, consider equation (28), that we rewrite as follows

$$\alpha_0\theta_{00}\left[1 + (\alpha_1 - \alpha_{00})/4\right]^2 - \alpha_{00}\theta_0 = 0. \tag{30}$$

We totally differentiate (30) to get

$$\frac{\hat{d}\theta_0}{\hat{d}\theta_{00}} = -\frac{\alpha_0 A^2 + \frac{A}{2}\theta_{00}\alpha_0\left[\alpha_1'\left(d\theta_1/d\theta_{00}\right) - \alpha_{00}'\right] - \alpha_{00}'\theta_0}{\alpha_0'\theta_{00}A^2 - \alpha_{00}}. \tag{31}$$

The denominator at the right-hand side of (31) can be rewritten as $\theta_{00}\left(\alpha_0' A^2 - \alpha_{00}/\theta_{00}\right) = \theta_{00}\left(\alpha_0' A^2 - \eta_{00}\right)$. Using (28), this expression can be further rewritten as $\theta_{00}A^2\left(\alpha_0' - \alpha_0/\theta_0\right)$ which is negative given our assumptions on the function $\alpha_i\left(\theta_i\right)$. Thus, to prove that $\hat{d}\theta_0/\hat{d}\theta_{00} > 0$, we need to show that the numerator at the right-hand side of (31) is positive. To compute $d\theta_1/d\theta_{00}$, we first rewrite (25) as

$$\frac{\alpha_1}{\theta_1} - \frac{\alpha_{00}x}{\theta_{00}} = 0.$$

It follows that

$$\frac{d\theta_1}{d\theta_{00}} = \frac{\theta_1}{\theta_{00}}\frac{\frac{\alpha_{00}'\theta_{00}}{\alpha_{00}} - 1}{\frac{\alpha_1'\theta_1}{\alpha_1} - 1}. \tag{32}$$

We confirm that $d\theta_1/d\theta_{00} > 0$ since $\alpha_{00}'\theta_{00} - \alpha_{00} < 0$ and $\alpha_1'\theta_1 - \alpha_1 < 0$. Using (25) and (32), the numerator at the right-hand side of (31) becomes

$$\frac{A}{2}\alpha_0\left(\alpha_1'\theta_1\frac{1 - \frac{\alpha_{00}'\theta_{00}}{\alpha_{00}}}{1 - \frac{\alpha_1'\theta_1}{\alpha_1}} - \theta_{00}\alpha_{00}'\right) + \alpha_0 A^2 - \alpha_{00}'\theta_0. \tag{33}$$

Consider the last term in (33), $\alpha_0 A^2 - \alpha_{00}'\theta_0$. Use (28) to rewrite it as $\theta_0\left(\alpha_{00} - \theta_{00}\alpha_{00}'\right)/\theta_{00}$ which is positive. Consider the term in brackets. Notice that $\alpha_i'\theta_i/\alpha_i < 1$ always holds since α_i is concave. If $\alpha_i'\theta_i/\alpha_i$ is weakly increasing in θ_i, it follows that $1 - \alpha_{00}'\theta_{00}/\alpha_{00} \geq 1 - \alpha_1'\theta_1/\alpha_1$, and $\alpha_{00}'\theta_{00} < \alpha_1'\theta_1$ since $\alpha_{00} < \alpha_1$. In this case the term in brackets is unambiguously positive. If $\alpha_i'\theta_i/\alpha_i$ is instead decreasing in θ_i, notice that, for $x \to 1$, $\theta_{00} \to \theta_1$ and hence the term is brackets tends to zero. Therefore, since α_i is a continuous function, if x is sufficiently high the overall expression in (33) is positive. We conclude that for x sufficiently high the numerator in (31) is always positive and therefore $\hat{d}\theta_0/\hat{d}\theta_{00} > 0$.

Since $d\theta_0/d\theta_{00} < 0$ while $\hat{d}\theta_0/\hat{d}\theta_{00} > 0$, we need to show that the intercept value of θ_0 that solves (27) for $\theta_{00} \to 0$ is greater than the intercept value of θ_0 that solves (28) for $\theta_{00} \to 0$. This will ensure that the curves defined in (27) and (28) intersect for some

$\theta_{00}, \theta_0 > 0$. Notice that when $\theta_{00} = 0$ then $\alpha_{00} = 0$ and $\eta_{00} = 1$, which given (25) implies that $\eta_1 = x < 1$ and hence $\alpha_1 > 0$.

When $\theta_{00} = 0$, (27) gives

$$\alpha_0 (\theta_0) = \frac{4k/x}{\left\{1 + \left[\alpha_1 \left(\eta_1^{-1} (x)\right)\right] /4\right\}^2}, \tag{34}$$

while (28) gives

$$\eta_0 (\theta_0) = \frac{1}{\left\{1 + \left[\alpha_1 \left(\eta_1^{-1} (x)\right)\right] /4\right\}^2}. \tag{35}$$

Consider three cases regarding the value of x and k; i.e., $x = 4k$, $x > 4k$ and $2k \leq x < 4k$ (notice that any equilibrium requires $x \geq 2k$). First consider the right-hand sides in (34) and (35) for $4k/x = 1$ and the maximum value of x; i.e., $x \to 1$. In this case, when $\theta_{00} = 0$, it follows that $\eta_{00} = \eta_1 = 1$, so that $\theta_1 = \theta_{00} = 0$ and thus $\alpha_1 = 0$. Hence the right-hand side in (34) becomes equal to 1, which implies that the value of θ_0 that solves (34) is $\theta_0 \to \infty$. Similarly, the right-hand side in (35) becomes equal to 1 which implies that the value of θ_0 that solves (35) is $\theta_0 = 0$. Therefore, for the extreme case $x \to 1$, the intercept value of θ_0 that solves (27) is greater than the intercept value of θ_0 that solves (28) for $\theta_{00} \to 0$. It follows that for $x < 1$ sufficiently high, when $\theta_{00} = 0$ then η_1 satisfies $\eta_1 < \eta_{00} = 1$ and is sufficiently high, hence θ_1 is sufficiently low, and α_1 is sufficiently low as well. Therefore the denominators in the right-hand sides of (34) and (35) are relatively small, and so the right-hand sides in (34) and (35) are relatively big, which ensures that the value of θ_0 that solves (34) (or (27) for $\theta_{00} = 0$) is greater than the value of θ_0 that solves (35) (or (28) for $\theta_{00} = 0$). We deduce that there is some $\hat{x}$ such that if $x \geq \hat{x}$ then the value of θ_0 that solves (34) is higher than the value of θ_0 that solves (35), and a unique solution for $\theta_0, \theta_{00} > 0$ exists. Second, consider the case $4k < x$. For $x \to 1$, the value of θ_0 that solves (35) is $\theta_0 = 0$ as in the previous case, while the value of θ_0 that solves (34) is $\theta_0 > 0$. Therefore, similarly as for $4k = x$, there is some $\hat{x}$ such that if $x \geq \hat{x}$ then the value of θ_0 that solves (34) is higher than the value of θ_0 that solves (35), and a unique solution $\theta_0, \theta_{00} >$ exists. Finally, consider the case $4k > x \geq 2k$. In this case, there is no value of θ_0 that yields $\alpha_0 \leq 1$ for $x \to 1$ and $\theta_{00} = 0$ in (35): k is big relative to x, so even if $x = 1$ (and also if $x < 1$), there is no θ_0 that makes sellers enter if $\theta_{00} = 0$. Hence there must be some $\theta_{00} > 0$ such that $\alpha_0 = 1$ (and so $\theta_0 \to \infty$) and the free-entry condition (27) holds. In the buyers' indifference condition (28), however, when $\theta_0 \to \infty$ we have that $\eta_0 = 0$ and hence it must be that θ_{00} also approaches infinity. Therefore the two curves defined by (27) and (28) must intersect for $\theta_0, \theta_{00} > 0$, and hence an equilibrium solution $\theta_0, \theta_{00} > 0$ exists.

In what follows we show that an increase in k decreases buyers' expected utility (sellers' free entry condition ensures that sellers' expected utility is always zero). Totally differentiate (27) with respect to k

$$\frac{4}{x} = \alpha_0' A^2 \frac{d\theta_0}{dk} + \left[\alpha_0 \alpha_1' \frac{A}{2} \frac{d\theta_1}{d\theta_{00}} + \alpha_{00}' \left(1 - \alpha_0 \frac{A}{2}\right)\right] \frac{d\theta_{00}}{dk} \tag{36}$$

where again $A = 1 + (\alpha_1 - \alpha_{00})/4$. Equation (36) shows that at least one of the two derivatives $d\theta_0/dk$ and $d\theta_{00}/dk$ must be positive.

Rewrite (26) as follows

$$\alpha_0 \theta_{00} \left(1 + \frac{\alpha_1 - \alpha_{00}}{4}\right)^2 - \alpha_{00} \theta_0 = 0$$

Totally differentiate with respect to k to obtain

$$\left(\alpha_0' \theta_{00} A^2 - \alpha_{00}\right) \frac{d\theta_0}{dk} + \left[\alpha_0 A^2 + \alpha_0 \theta_{00} \frac{A}{2} \left(\alpha_1' \frac{d\theta_1}{d\theta_{00}} - \alpha_{00}'\right) - \theta_0 \alpha_{00}'\right] \frac{d\theta_{00}}{dk} = 0 \tag{37}$$

Combining (36) and (37), we obtain

$$\frac{d\theta_{00}}{dk} = \frac{\frac{4}{x}\left(\frac{\alpha_0}{\theta_0 \alpha_0'} - 1\right)}{\frac{(\alpha_0)^2 \left(1 + \frac{\alpha_1 - \alpha_{00}}{4}\right)}{\theta_0 \alpha_0'} \frac{\alpha_1' \frac{d\theta_1}{d\theta_{00}} - \alpha_{00}'}{2} + \left(\frac{\alpha_0}{\theta_0 \alpha_0'} - 1\right) \alpha_{00}' + \frac{\theta_0}{\theta_{00}} \left(\frac{\alpha_{00}}{\theta_{00}} - \alpha_{00}'\right)} \tag{38}$$

On the right-hand side in (38), the numerator is positive, thus if the denominator is also positive then $d\theta_{00}/dk > 0$. Notice that the denominator is positive if

$$\theta_{00} \frac{(\alpha_0)^2 A}{\theta_0 \alpha_0'} \frac{\alpha_1' \frac{d\theta_1}{d\theta_{00}} - \alpha_{00}'}{2} + \theta_{00} \left(\frac{\alpha_0}{\theta_0 \alpha_0'} - 1\right) \alpha_{00}' + \theta_0 \left(\frac{\alpha_{00}}{\theta_{00}} - \alpha_{00}'\right) > 0 \tag{39}$$

To show that (39) holds, recall that existence and uniqueness of the equilibrium required that the overall term in (33) or equivalently the numerator on the right-hand side of (31) to be positive. By subtracting the numerator on the right-hand side of (31) from the term on the left-hand side of (39) and rearranging, we obtain the following condition for the denominator on the right-hand side in (38) to be positive

$$\alpha_0 \frac{A}{2} \alpha_1' \frac{d\theta_1}{d\theta_{00}} + \alpha_{00}' \left(1 - \alpha_0 \frac{A}{2}\right) > 0 \tag{40}$$

which holds since $\alpha_0 A/2 < 1$. It follows that in the defined equilibrium $d\theta_{00}/dk > 0$, which implies that buyers' expected utility of searching in market 00 decreases with k: since θ_{00} increases η_{00} decreases with k, while $p_{00} = x/2$ does not change with k. Since buyers' expected utility must be the same across markets in equilibrium, buyers' expected utility of searching in the three markets decreases with k.

Appendix B. Regressions

Table 9: Rome. Dependent variable: average daily rate (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.216***	0.215***	0.211***	0.207***	0.198***	0.198***
	(0.00899)	(0.00898)	(0.00902)	(0.00902)	(0.00885)	(0.00887)
Max Guests	0.0786***	0.0788***	0.0804***	0.0787***	0.0758***	0.0758***
	(0.00275)	(0.00275)	(0.00275)	(0.00277)	(0.00271)	(0.00271)
superhost2		0.0378**	0.00400	0.00197	0.0254*	0.0256*
		(0.0115)	(0.0120)	(0.0120)	(0.0118)	(0.0118)
Overall Rating			0.124***	0.117***	0.125***	0.125***
			(0.0117)	(0.0117)	(0.0115)	(0.0115)
Number of Photos				0.00134***	0.00176***	0.00174***
				(0.000297)	(0.000292)	(0.000294)
Number of Reviews					-0.00139***	-0.00139***
					(0.000103)	(0.000104)
instantbook						0.00297
						(0.00789)
Constant	4.022***	4.019***	3.549***	3.553***	3.579***	3.578***
	(0.0669)	(0.0668)	(0.0750)	(0.0748)	(0.0732)	(0.0733)
Observations	4169	4169	3978	3971	3971	3971
Adjusted R^2	0.525	0.526	0.544	0.546	0.566	0.566

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Rome. Dependent variable: occupancy rate ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.0191**	0.0197**	0.0195**	0.0171**	0.0178**	0.0191**
	(0.00663)	(0.00648)	(0.00653)	(0.00647)	(0.00591)	(0.00592)
Max Guests	0.00465*	0.00606**	0.00819***	0.00631**	0.00550**	0.00548**
	(0.00208)	(0.00203)	(0.00205)	(0.00205)	(0.00187)	(0.00187)
logadr	-0.182***	-0.190***	-0.209***	-0.215***	-0.157***	-0.157***
	(0.0107)	(0.0105)	(0.0108)	(0.0107)	(0.0100)	(0.0100)
superhost2		0.111***	0.0765***	0.0726***	0.0412***	0.0423***
		(0.00778)	(0.00812)	(0.00806)	(0.00745)	(0.00745)
Overall Rating			0.0953***	0.0888***	0.0710***	0.0709***
			(0.00803)	(0.00800)	(0.00734)	(0.00733)
Number of Photos				0.00184***	0.00120***	0.00113***
				(0.000200)	(0.000184)	(0.000185)
Number of Reviews					0.00185***	0.00183***
					(0.0000663)	(0.0000666)
instantbook						0.0170***
						(0.00496)
Constant	1.419***	1.441***	1.116***	1.128***	0.885***	0.878***
	(0.0633)	(0.0618)	(0.0636)	(0.0632)	(0.0584)	(0.0583)
Observations	4169	4169	3978	3971	3971	3971
Adjusted R^2	0.091	0.133	0.161	0.178	0.313	0.315

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Rome. Dependent variable: annual revenue (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.140***	0.137***	0.125***	0.109***	0.174***	0.179***
	(0.0217)	(0.0212)	(0.0216)	(0.0214)	(0.0172)	(0.0172)
Max Guests	0.0500***	0.0529***	0.0548***	0.0459***	0.0652***	0.0651***
	(0.00664)	(0.00649)	(0.00657)	(0.00655)	(0.00526)	(0.00525)
superhost2		0.387***	0.277***	0.265***	0.107***	0.111***
		(0.0272)	(0.0286)	(0.0283)	(0.0229)	(0.0229)
Overall Rating			0.273***	0.237***	0.181***	0.181***
			(0.0279)	(0.0278)	(0.0223)	(0.0222)
Number of Photos				0.00684***	0.00403***	0.00376***
				(0.000702)	(0.000566)	(0.000568)
Number of Reviews					0.00939***	0.00931***
					(0.000200)	(0.000200)
instantbook						0.0617***
						(0.0153)
Constant	9.009***	8.977***	7.919***	7.930***	7.752***	7.724***
	(0.162)	(0.158)	(0.179)	(0.177)	(0.142)	(0.142)
Observations	4169	4169	3978	3971	3971	3971
Adjusted R^2	0.081	0.123	0.142	0.160	0.461	0.463

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: San Francisco. Dependent variable: average daily rate (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.296***	0.293***	0.283***	0.275***	0.267***	0.267***
	(0.0191)	(0.0189)	(0.0186)	(0.0187)	(0.0186)	(0.0187)
Max Guests	0.0877***	0.0897***	0.0933***	0.0887***	0.0880***	0.0884***
	(0.00400)	(0.00397)	(0.00390)	(0.00399)	(0.00396)	(0.00400)
superhost2		0.107***	0.0641***	0.0516**	0.0660***	0.0660***
		(0.0167)	(0.0169)	(0.0168)	(0.0169)	(0.0169)
Overall Rating			0.214***	0.210***	0.210***	0.208***
			(0.0208)	(0.0209)	(0.0208)	(0.0210)
Number of Photos				0.00383***	0.00449***	0.00450***
				(0.000591)	(0.000601)	(0.000601)
Number of Reviews					-0.000840***	-0.000829***
					(0.000160)	(0.000161)
instantbook						-0.0112
						(0.0172)
Constant	4.739***	4.718***	3.724***	3.700***	3.730***	3.740***
	(0.0228)	(0.0228)	(0.0992)	(0.0995)	(0.0990)	(0.100)
Observations	2069	2069	2029	1995	1995	1995
Adjusted R^2	0.583	0.591	0.614	0.620	0.625	0.625

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood and city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: San Francisco. Dependent variable: occupancy rate ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.0159	0.0179	0.0188	0.0175	0.0240*	0.0259*
	(0.0123)	(0.0122)	(0.0122)	(0.0124)	(0.0119)	(0.0119)
Max Guests	0.00508	0.00762**	0.0116***	0.0100***	0.00913***	0.00748**
	(0.00271)	(0.00270)	(0.00276)	(0.00281)	(0.00270)	(0.00270)
logadr	-0.153***	-0.167***	-0.187***	-0.193***	-0.172***	-0.171***
	(0.0136)	(0.0136)	(0.0140)	(0.0143)	(0.0138)	(0.0137)
superhost2		0.0726***	0.0576***	0.0536***	0.0306**	0.0304**
		(0.0103)	(0.0105)	(0.0106)	(0.0103)	(0.0103)
Overall Rating			0.0729***	0.0729***	0.0692***	0.0780***
			(0.0133)	(0.0135)	(0.0129)	(0.0130)
Number of Photos				0.00158***	0.000489	0.000466
				(0.000376)	(0.000370)	(0.000368)
Number of Reviews					0.00128***	0.00123***
					(0.0000981)	(0.0000980)
instantbook						0.0500***
						(0.0104)
Constant	1.474***	1.524***	1.281***	1.295***	1.170***	1.121***
	(0.0659)	(0.0655)	(0.0809)	(0.0820)	(0.0792)	(0.0794)
Observations	2069	2069	2029	1995	1995	1995
Adjusted R^2	0.097	0.119	0.132	0.138	0.207	0.216

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood and city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: San Francisco. Dependent variable: annual revenue (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.201***	0.190***	0.180***	0.151**	0.242***	0.247***
	(0.0479)	(0.0469)	(0.0476)	(0.0476)	(0.0406)	(0.0406)
Max Guests	0.0828***	0.0901***	0.0966***	0.0831***	0.0904***	0.0874***
	(0.0100)	(0.00985)	(0.00999)	(0.0101)	(0.00863)	(0.00870)
superhost2		0.390***	0.346***	0.296***	0.133***	0.133***
		(0.0415)	(0.0432)	(0.0428)	(0.0369)	(0.0369)
Overall Rating			0.178***	0.157**	0.163***	0.180***
			(0.0533)	(0.0532)	(0.0453)	(0.0457)
Number of Photos				0.0133***	0.00585***	0.00581***
				(0.00150)	(0.00131)	(0.00131)
Number of Reviews					0.00950***	0.00941***
					(0.000349)	(0.000351)
instantbook						0.0934*
						(0.0375)
Constant	9.515***	9.440***	8.614***	8.573***	8.227***	8.142***
	(0.0571)	(0.0565)	(0.254)	(0.253)	(0.216)	(0.218)
Observations	2069	2069	2029	1995	1995	1995
Adjusted R^2	0.161	0.196	0.205	0.238	0.449	0.450

Standard errors in parentheses

These are OLS regressions. Each regression controls for neighborhood and city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Miami. Dependent variable: average daily rate (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.345***	0.346***	0.352***	0.342***	0.323***	0.324***
	(0.0137)	(0.0137)	(0.0136)	(0.0134)	(0.0130)	(0.0130)
Max Guests	0.0947***	0.0953***	0.0977***	0.0933***	0.0901***	0.0897***
	(0.00354)	(0.00355)	(0.00351)	(0.00351)	(0.00339)	(0.00340)
superhost2		0.0456**	0.0329	0.0202	0.0562***	0.0561***
		(0.0176)	(0.0175)	(0.0173)	(0.0169)	(0.0169)
Overall Rating			0.0960***	0.0886***	0.101***	0.101***
			(0.0128)	(0.0127)	(0.0122)	(0.0122)
Number of Photos				0.00356***	0.00433***	0.00430***
				(0.000381)	(0.000370)	(0.000370)
Number of Reviews					-0.00316***	-0.00319***
					(0.000183)	(0.000183)
instantbook						0.0187
						(0.0111)
Constant	3.968***	3.960***	3.474***	3.488***	3.509***	3.503***
	(0.0428)	(0.0429)	(0.0731)	(0.0723)	(0.0697)	(0.0698)
Observations	4261	4261	4034	4032	4032	4032
Adjusted R^2	0.454	0.455	0.492	0.502	0.537	0.537

Standard errors in parentheses

These are OLS regressions. Each regression controls for city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16: Miami. Dependent variable: occupancy rate ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.0501***	0.0517***	0.0509***	0.0501***	0.0478***	0.0522***
	(0.00697)	(0.00690)	(0.00710)	(0.00704)	(0.00681)	(0.00674)
Max Guests	-0.00152	-0.000217	0.00366*	0.00275	0.00114	0.000237
	(0.00181)	(0.00180)	(0.00186)	(0.00185)	(0.00179)	(0.00177)
logadr	-0.203***	-0.206***	-0.222***	-0.231***	-0.197***	-0.199***
	(0.00727)	(0.00721)	(0.00765)	(0.00768)	(0.00769)	(0.00760)
superhost2		0.0784***	0.0595***	0.0547***	0.0363***	0.0363***
		(0.00827)	(0.00848)	(0.00843)	(0.00822)	(0.00812)
Overall Rating			0.0597***	0.0574***	0.0483***	0.0498***
			(0.00623)	(0.00619)	(0.00600)	(0.00593)
Number of Photos				0.00150***	0.000995***	0.000904***
				(0.000187)	(0.000183)	(0.000181)
Number of Reviews					0.00156***	0.00148***
					(0.0000923)	(0.0000915)
instantbook						0.0532***
						(0.00536)
Constant	1.488***	1.485***	1.260***	1.298***	1.169***	1.160***
	(0.0353)	(0.0349)	(0.0443)	(0.0442)	(0.0434)	(0.0429)
Observations	4261	4261	4034	4032	4032	4032
Adjusted R^2	0.205	0.222	0.239	0.251	0.300	0.317

Standard errors in parentheses

These are OLS regressions. Each regression controls for city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17: Miami. Dependent variable: annual revenue (log) ($\geq$ 10 bkngs)

	(1)	(2)	(3)	(4)	(5)	(6)
Bathrooms	0.294***	0.296***	0.279***	0.258***	0.313***	0.317***
	(0.0232)	(0.0230)	(0.0236)	(0.0232)	(0.0209)	(0.0209)
Max Guests	0.0489***	0.0527***	0.0608***	0.0513***	0.0606***	0.0597***
	(0.00598)	(0.00594)	(0.00611)	(0.00605)	(0.00544)	(0.00545)
superhost2		0.280***	0.219***	0.191***	0.0873**	0.0872**
		(0.0295)	(0.0304)	(0.0299)	(0.0271)	(0.0270)
Overall Rating			0.154***	0.138***	0.102***	0.103***
			(0.0222)	(0.0218)	(0.0196)	(0.0196)
Number of Photos				0.00765***	0.00542***	0.00533***
				(0.000657)	(0.000594)	(0.000594)
Number of Reviews					0.00915***	0.00909***
					(0.000293)	(0.000294)
instantbook						0.0456*
						(0.0179)
Constant	8.687***	8.638***	7.909***	7.940***	7.880***	7.867***
	(0.0724)	(0.0718)	(0.127)	(0.125)	(0.112)	(0.112)
Observations	4261	4261	4034	4032	4032	4032
Adjusted R^2	0.156	0.174	0.190	0.216	0.369	0.369

Standard errors in parentheses

These are OLS regressions. Each regression controls for city.

Source: Authors' calculations from AIRDNA data

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$