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 then use 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 more bookings 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.

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 their price..... (**NEED TO DESCRIBE THE RESULTS**). 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 in order to attract buyers and if a star is received, it then charges a higher price. 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.

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 more bookings and more revenue in each of the four cities we examine.

The paper is organized as follows. Section 1 contains a brief literature review. Section 2 describes the search enviroment and the seller's pricing decisions. Section 3 describes the welfare benefits of a rating system. Section 4 describes the data we use. Section 5 presents the empirical findings. Section 6 concludes.

2 Literature Review

To do

3 Basic features of Airbnb data

We analyse data from Airbnb, an online platform that allows hosts and guests to be matched. Hosts can post their listings, and guests can search for rentals that best suit their preferences.[1] 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.

The key aspect of the data that we explore is the role of the superhost status that hosts may acquire through Airbnb. 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. 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

[1] Most of the data was colllected in July 2016. See the appendix for details.

appears on the superhost' profile and listing pages and precisely certifies that the host has complied with the superhost' requirements.[2]

Table 1 presents the average daily rate, the number of bookings in the last twelve months and the annual revenue sorted by superhost status for Amsterdam, Rome, Miami and San Francisco. 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.

[2]The other benefits are ...

TABLE 1.

Differences in averages between superhosts and non-superhosts - All rentals

	Superhost	Non-superhost	Diff.
	(1)	(2)	(1)-(2)
Amsterdam			
average daily rate	$167,4$	$159,7$	$7,7^{***}$
occupancy rate	$0,67$	$0,59$	$0,08^{***}$
number of bookings	$29,4$	$14,9$	$14,5^{***}$
annual revenue	16492	8394	8098^{***}
obs.	1368	14520	
Rome			
average daily rate	$125,3$	$123,5$	$1,8$
occupancy rate	$0,57$	$0,42$	$0,15^{***}$
number of bookings	$31,6$	$13,6$	$18,0^{***}$
annual revenue	15987	7422	8565^{***}
obs.	1403	15609	
Miami			
average daily rate	$175,2$	$226,3$	$-51,1^{***}$
occupancy rate	$0,56$	$0,44$	$0,12^{***}$
number of bookings	$19,1$	$8,1$	$11,0^{***}$
annual revenue	14352	10086	4266^{***}
obs.	1718	18220	
San Francisco			
average daily rate	$188,3$	$208,4$	$-20,1^{***}$
occupancy rate	$0,63$	$0,56$	$0,07^{***}$
number of bookings	$27,9$	$9,7$	$18,2^{***}$
annual revenue	21345	10780	10565^{***}
obs.	2767	15950	

Source: Authors' calculations from AIRDNA data.

TABLE 2.

Differences in averages between superhosts and non-superhosts - Apartments, Entire

	Superhost	Non-superhost	Diff.
	(1)	(2)	(1)-(2)
Amsterdam			
average daily rate	$175,3$	$168,0$	$7,3^{**}$
occupancy rate	$0,67$	$0,60$	$0,07^{***}$
number of bookings	$24,3$	$10,2$	$14,1^{***}$
annual revenue	16487	8510	7977^{***}
obs.			
Rome			
average daily rate	$138,2$	$142,4$	$-4,2^{**}$
occupancy rate	$0,60$	$0,44$	$0,16^{***}$
number of bookings	$33,8$	$14,8$	$19,0^{***}$
annual revenue	18555	9150	9405^{***}
obs.			
Miami			
average daily rate	$211,2$	$231,1$	$-19,9^{**}$
occupancy rate	$0,59$	$0,46$	$0,13^{***}$
number of bookings	$19,0$	$8,7$	$10,3^{***}$
annual revenue	19471	11654	7817^{***}
obs.			
San Francisco			
average daily rate	$229,6$	$237,7$	$8,1$
occupancy rate	$0,70$	$0,60$	$0,10^{***}$
number of bookings	$28,7$	$10,3$	$18,4^{***}$
annual revenue	29347	14425	14922^{***}
obs.			

Source: Authors' calculations from AIRDNA data.

TABLE 3.

Differences in averages superhosts and non-superhosts - Apt, Ent., 1 bed, 1 bath, max 2 g.

	Superhost	Non-superhost	Diff.
	(1)	(2)	(1)-(2)
Amsterdam			
average daily rate	147,5	135,4	
occupancy rate	0,68	0,62	
number of bookings	24,5	10,9	
annual revenue	14253	6061	
obs.	391	3842	
Rome			
average daily rate	108,7	100,5	
occupancy rate	0,63	0,48	
number of bookings	36,2	14,4	
annual revenue	16529	7662	
obs.	119	912	
Miami			
average daily rate	146,1	157,4	-11,3
occupancy rate	0,59	0,48	
number of bookings	18,7	7,3	
annual revenue	15776	8182	
obs.	73	820	
San Francisco			
average daily rate	175,6	180,0	-4,4
occupancy rate	0,71	0,60	
number of bookings	32,4	9,4	
annual revenue	27147	11290	
obs.	186	1151	

Source: Authors' calculations from AIRDNA data.

Notice that, when considering all rentals, in Rome and San Francisco non-superhosts charge a higher average daily rate than superhosts. However, this is not the case anymore once we control for the physical characteristics of the rentals. In all cities most rentals have the following features: in terms of listing type they are entire homes or apartments, in terms of property type they are apartments, they have 1 bedroom and 1 bathroom and are suited for at most 2 guests (the breaking down of these categories for each city is presented in the appendix). Table 3 displays the data for the subset of rentals that satisfy these features. In the four cities, superhosts either charge a higher average rate than non-superhosts or the difference in the rates set by superhosts and non-superhosts is not statistically significant. Since superhosts reach a higher number of bookings than non-superhosts, this results in higher annual revenues for these hosts.[3]

The overall rating (average number of stars) and the average number of reviews received by superhosts and non-superhosts are presented in Table 4 (for the same subset of rentals as in Table 3). Consistently with the superhosts' requirements, the average values of these two variables are systematically higher for superhosts. Table 4 also presents the average response time and the fraction of hosts with a response rate of 100%. As expected, superhosts are more prone to have a 100% response rate and respond faster to their guests than non-superhosts.

[3] Occupancy rate... Appendix: neighborhoods or areas.

TABLE 4.

Differences in averages between superhosts and non-superhosts - Apt, Ent., 1 bed, 1 bath, ma

	Superhost	Non-superhost	Diff.
	(1)	(2)	(1)-(2)
Amsterdam			
overall rating	4,85	4,68	
nr. reviews	33,1	12,0	
resp. rate=100% (%)	85,0	63,8	
resp. time (min)	205,1	317,6	
obs.			
Rome			
overall rating	4,83	4,52	
nr. reviews	50,6	18,5	
resp. rate=100% (%)	92,3	74,8	
resp. time (min)	81,2	189,4	
obs.			
Miami			
overall rating	4,87	4,56	
nr. reviews	23,0	8,46	
resp. rate=100% (%)	86,3	51,5	
resp. time (min)	71,9	265,3	
obs.			
San Francisco			
overall rating	4,86	4,68	
nr. reviews	49,9	12,6	
resp. rate=100% (%)	88,7	62,3	
resp. time (min)	126,5	307,6	
obs.			

Source: Authors' calculations from AIRDNA data.

In the next section, we present a model that is consistent with this data.

In this model, the equilibrium is such that sellers that acquire a good rating charge high prices and have a high probability of materialising a sale, compared to unrated sellers. As mentioned in the literature review, the interest of such a model...

4 A search model with heterogenous quality

4.1 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 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 their 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 $\alpha_i = M(s_i, b_i)/s_i$. Let θ_i denote tightness in market i, where, for $i \in \{0, 1, 00\}$, 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,

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

4.2 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).

4.2.1 Buyers' acceptance decisions

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

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

and in the rated market the expected utility is

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

Buyers direct their search towards the market that yields the highest expected utility.

4.2.2 Sellers' price posting decisions

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

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

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

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

In each state, the sellers chose price p_i in order to maximize their lifetime utility. That is they chose p_i such that the right hand-sides of the value functions (8)-(10) are maximized. The first-order conditions are

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

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

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

Replacing p_1 in (9) and p_{00} in (10), we can get 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

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

If the arrival rates for old sellers are equal; i.e., $\alpha_1 = \alpha_{00}$, then $p_0 = p_{00}$. However, if $\alpha_1 > \alpha_{00}$, then $p_0 < p_{00} = x/2$.

4.2.3 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:

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

Then free entry implies

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

Use (11) to replace p_0 to get:

$$k = \alpha_0 x \left[1 + \left(\alpha_1 - \alpha_{00}\right)/4\right]^2/4 + x\alpha_{00}/4 \qquad (16)$$

4.3 Rating Equilibrium

Buyers observe prices (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 need to be indifferent between the three options. From (6) and (7), they are indifferent if

$$\eta_0 \int_{p_0/x}^{1} \left(x\varepsilon - p_0\right) d\varepsilon = \eta_{00} \int_{p_{00}/x}^{1} \left(x\varepsilon - p_{00}\right) d\varepsilon = \eta_1 \int_{p_1}^{1} \left(\varepsilon - p_1\right) d\varepsilon \qquad (17)$$

Using (12) and (13), the second equality of (17) simplifies to

$$\eta_1 = \eta_{00} x \qquad (18)$$

Thus, for a buyer, the probability of a match is smaller in the rated market than in the unrated one if $x < 1$.

Lemma 1 *In any equilibrium with $x < 1$, $p_1 > p_{00} > p_0$. For $x = 1$ we have $p_1 = p_{00} = p_0$.*

Proof. From (18), 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 $p_0 < p_{00} = x/2 < p_1 = 1/2$. If $x = 1$, then from (18), we have $\eta_1 = \eta_{00}$ and so $\alpha_1 = \alpha_{00}$ implying that $p_0 = p_{00} = p_1 = 1/2$. ■

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

In the following, we show how to compute the equilibrium. The first equality of (17) 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 follows:

$$\eta_0 \left[1 + \left(\alpha_1 - \alpha_{00}\right)/4\right]^2 = \eta_{00} \tag{19}$$

Proposition 2 *A rating equilibrium are prices p_i and tightnesses θ_i, $i \in \{0, 1, 00\}$, that solve (12), (13), and (15)-(19). A rating equilibrium exists and is unique.*

Proof. See Appendix. ■

4.3.1 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 he can go to the unrated market even if he is a low-quality seller. Accordingly, his life-time utility is

$$\tilde{V}_0 = \alpha_0 \int_{p_0/x}^{1} d\varepsilon \left[p_0 \right] + V_{00}$$

while a seller that hides this information and sells in the unrated market gets

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

The equilibrium expected payoff is

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

Thus we need

$$\begin{aligned} \left(1 - p_0/x\right) &> p_0/x \\ p_0 &< x/2 \\ \frac{x\left[1 - \left(\alpha_1 - \alpha_{00}\right)/4\right]}{2} &< x/2 \\ 1 - \left(\alpha_1 - \alpha_{00}\right)/4 &< 1 \\ \alpha_{00} &< \alpha_1 \end{aligned}$$

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.

4.4 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 conditions 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 denoted by V_0 and V_1 for old and young:

$$V_0 = \alpha_0 \int_{p_0/x}^{1} d\varepsilon p_0 + V_1$$

$$V_1 = \alpha_1 \int_{p_1/x}^{1} d\varepsilon p_1 = \alpha_1 x \left(1 - p_1/x\right) p_1/x$$

We can simplify the V_0 as follows

$$V_0 = \alpha_0 x \left(1 - p_0/x\right) p_0/x + \alpha_1 x \left(1 - p_1/x\right) p_1/x$$

The first-order conditions for p_0 and p_1 are

$$p_0/x = p_1/x = 0.5$$

The free entry condition implies that

$$k = V_0 = \alpha_0 x/4 + \alpha_1 x/4$$

Buyers have to be indifferent between the two markets

$$\eta_0 \int_{p_0/x}^{1} (x\varepsilon - p_0)\, d\varepsilon = \eta_1 \int_{p_1/x}^{1} (x\varepsilon - p_1)\, d\varepsilon$$

Since the two prices are the same, the probabilities are the same and so tightness is the same. This implies that

$$\alpha_0 = \alpha_1 = 2k/x.$$

The buyer's expected utility then is

$$\begin{aligned}
\eta_0 \int_{p_0/x}^{1} (x\varepsilon - p_0)\, d\varepsilon &= \eta_0 \left[x \int_{p_0/x}^{1} \varepsilon d\varepsilon - \int_{p_0/x}^{1} p_0 d\varepsilon \right] \\
&= \eta_0 x \left\{ \left[\frac{1}{2} - \frac{(p_0/x)^2}{2} \right] - \left[p_0/x - (p_0/x)^2 \right] \right\} \\
&= \eta_0 x \left\{ \frac{1}{2} + \frac{(p_0/x)^2}{2} - p_0/x \right\} \\
&= \eta_0 x/8 = \eta_1 x/8
\end{aligned}$$

With ratings, the buyer's expected utility is

$$\begin{aligned}
\eta_1 \int_{p_1}^{1} (\varepsilon - p_1)\, d\varepsilon &= \eta_1 \left(\frac{1}{2} - \frac{(p_1)^2}{2} - (1 - p_1)\, p_1 \right) \\
&= \eta_1 \left(\frac{1}{2} + \frac{(p_1)^2}{2} - p_1 \right) \\
&= \eta_1/8
\end{aligned}$$

since $p_1 = 1/2$. The buyers are better off with ratings if

$$\eta_1 = \eta_{00}x > \bar{\eta}_1 x = \bar{\eta}_0 x$$

5 Characteristics of superhosts

In this section, we present additional hosts' characteristics that are available in our data set and that might be related to the observed pattern of prices and sales across hosts. In order to get a better grasp of the role played by the superhost label, we first investigate whether these patterns hold if the comparison between superhosts and non-superhosts is restricted to listings with more than 10 bookings per year. In Table 5, we present the average daily rate, the number of bookings in the last twelve months and the annual revenue by superhost status, for rentals that are entire apartments with one bedroom and one bathroom, that can host 2 guests at most, and with more than 10 bookings in the last twelve months. The patterns highlighted before hold for the four cities: both the average daily rates and the number of bookings are higher for superhosts, which results in higher annual revenues for these hosts.[4]

[4]In Rome and San Francisco, the difference in average daily rates is however not significant: ... Appendix: neighborhoods or areas.

TABLE 5.

Diff. averages superhosts / non-superhosts - Apt, Ent, 1 bed, 1 bath, max 2 g., >10 bkgns.

	Superhost	Non-superhost	Diff.
	(1)	(2)	(1)-(2)
Amsterdam			
average daily rate	147.7	134.2	
occupancy rate	0.70	0.66	
number of bookings	29.6	24.0	
annual revenue	17075	12358	
obs.	304	1272	
Rome			
average daily rate	106.0	94.9	
occupancy rate	0.67	0.60	
number of bookings	45.9	35.9	
annual revenue	19009	13127	
obs.	101	436	
Miami			
average daily rate	140.2	130.0	
occupancy rate	0.65	0.62	
number of bookings	34.0	27.5	
annual revenue	21137	16257	
obs.	49	258	
San Francisco			
average daily rate	175.2	167.6	
occupancy rate	0.73	0.69	
number of bookings	40.3	29.4	
annual revenue	30805	22260	
obs.	158	432	

Source: Authors' calculations from AIRDNA data.

Second, we consider the type of cancellation policy (flexible, moderate or strict) set by hosts for their rentals. Table 6 displays the cancellation policy by superhost status for the rentals that are entire apartments with one bedroom and one bathroom that can host 2 guests. Superhosts follow a more strict cancellation policy than non-superhosts in the four cities considered. Notice that this means that the expected rates are even higher for superhosts than for non-superhosts once cancellation policy, and hence the probability that guests cancel, is taken into account.[5]

[5] Take later cancellation policy into account for calibration. Explore nr available days to show that average daily rate is sometimes lower for superhosts maybe because they rent on a more regular basis.

TABLE 6

Diff. averages superhosts / non-superhosts - Apt, Ent, 1 bed, 1 bath, max 2 g., >10 bkgns.

	Superhost	Non-superhost	Diff.
Cancellation policy (%)	(1)	(2)	(1)-(2)
Amsterdam			
Flexible	16.8	28.4	
Moderate	44.3	43.5	
Strict	38.9	28.1	
obs.			
Rome			
Flexible	13.1	27.3	
Moderate	33.9	35.3	
Strict	53.1	37.4	
obs.			
Miami			
Flexible	17.9	32.0	
Moderate	26.3	14.3	
Strict	55.8	51.1	
obs.			
San Francisco			
Flexible	11.3	33.1	
Moderate	42.7	32.2	
Strict	46.1	34.2	
obs.			

Source: Authors' calculations from AIRDNA data.

6 Calibration

The model can be used to predict unobservable variables, such as the fraction x of high-quality sellers and tightnesses θ_1, θ_0 and θ_{00} that prevail in the market with rated sellers, unrated sellers that are in their first period of life and unrated sellers that are in their second period of life, respectively. We calibrate the model using Airbnb data for the four cities Amsterdam, Rome, Miami and San Francisco. We consider rentals that are entire apartments, with 1 bedroom and 1 bathroom, for 2 guests maximum, and with more than 10 bookings.

For the baseline calibration, we use the three following targets observed in the data: the proportion of superhosts, the ratio of superhosts' average revenue to non-superhosts' average revenue, and the ratio of the superhosts' average rate to the non-superhosts' average rate.

The price p_1 set by rated sellers in the model corresponds to the average daily rate set by superhosts p_S; i.e, $p_1 = p_S$. Concerning the unrated sellers, we interpret the difference between the "young" and "old" unrated sellers as the fact that the former appear as "new" on the listings, which allows them to get their chance of being positively rated and hence attract more guests, whereas those who appear as older becoming superhost is unlikely because they already have a relatively bad rating (i.e., they do not have 5 stars in 80% of their reviews as required for the superhost status). In the data, it is not straightforward to distinguish new or "young" sellers, who did not get a high-quality label but may still get it (the unrated sellers in their first period of life in our model) and "old" sellers, those for whom getting this label is already unlikely (the unrated sellers in their second period of life). Buyers who search for rentals on Airbnb get some information that allows them to infer whether rentals have a long record or if they are relatively new on the listings (e.g., the number of reviews), but this is not visible in the dataset.[6][7] The rate observed in the data is an average across both types of sellers. We let p_{NS} denote the average daily rate set by non-superhosts and y denote the

[6]In our model, α_0 is the chance of making a sale, but in the data we interpret α_0 more broadly not only as the chance of making a sale, but also as the chance that guests leave reviews and that these reviews are sufficiently positive.

[7]explain why occupancy rate is not good indicator.

proportion of non-superhosts that are relatively "young" and correspond to the sellers in their first period of life in our model ($(1-y)$ is the proportion of non-superhosts that are relatively "old"). Then $p_{NS} = yp_0 + (1-y)\,p_{00}$; i.e., the observed price by non-superhosts is the weighted average of the price set by "young" non-superhosts and "old" non-superhosts.

In the model, average revenue by superhosts is $\alpha_1 p_1$. Similarly, average revenue by non-superhosts is $y\alpha_0 p_0 + (1-y)\,\alpha_{00} p_{00}$. Since the theoretical model predicts $p_1 > p_0, p_{00}$ and $\alpha_1 > \alpha_0, \alpha_{00}$, this implies that revenue by superhosts is higher than revenue by non-superhosts; i.e., $\alpha_1 p_1 > a_0 p_0, \alpha_{00} p_{00}$.

We assume that $\alpha_i\,(\theta_i)$ takes the functional form $\alpha_i = z\,(\theta_i)^c$. We normalise z and consider values of c consistent with $\alpha_i, \eta_i \leq 1$. Table 7 presents the parameters and the targets of the baseline calibration.

TABLE 7

Parameters	Targets				
		Amsterdam	Rome	Miami	San F.
x	fraction of superhosts	19.1%	18.8%	16.0%	26.8%
y	revenue: ratio superhosts to non-superhosts	1.38	1.45	1.30	1.38
k	daily rate: ratio superhosts to non-superhosts	1.14	1.16	1.11	1.14
z	set to 0.3				
c	set to 0.6				

Table 8 presents the calibrated parameters and variables for each city. In all cities, the proportion of high-quality sellers implied by the calibrated model lies between 0.86 and 0.9. The proportion of young unrated sellers takes the lowest value in Miami (0.07%) and the largest value in Rome (15.1%). The entry cost is largest for San Francisco (0.133) and smallest for Miami (0.078). Intuively, in San Francisco the largest entry costs are offset by the largest sales' probabilities α_0, α_{00} and α_1.

TABLE 8

Calibrated values				
	Amsterdam	Rome	Miami	San Francisco
x	0.878	0.861	0.900	0.878
y	0.046	0.151	0.007	0.020
k	0.095	0.095	0.078	0.133
θ_0	0.580	0.172	0.418	1.029
θ_{00}	0.548	0.167	0.403	0.952
θ_1	0.758	0.813	0.523	1.318
α_0	0.216	0.104	0.178	0.305
α_{00}	0.209	0.102	0.174	0.291
α_1	0.254	0.128	0.203	0.354

7 Extensions

The model basically shows that sellers want to be rated because it increases arrival rates and they can charge a higher price.

Where should we go with this?

We can have several extensions:

- Model without different qualities but with learning by doing. This will show that we also get a higher price in the second period but in contrast to the model above arrival rates will be lower.

- Model with infinitely lived sellers that can produce twice. This model is similar to the one presented here. The big difference is that even when we set $x = 0$, we get that the sellers charge a lower price in market one. This is because of discounting.

- More ideas?

References

Appendix

PROOF PROPOSITION 1

Proof. Equilibrium tightnesses solve (16)-(19), which we restate here for ease of reference:

$$k - \alpha_0 x \left[1 + (\alpha_1 - \alpha_{00})/4\right]^2/4 - x\alpha_{00}/4 = 0$$
$$\eta_1 - \eta_{00} x = 0$$
$$\eta_0 \left[1 + (\alpha_1 - \alpha_{00})/4\right]^2 - \eta_{00} = 0$$

From the second equation, 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} :

$$4k/x - \alpha_0(\theta_0)\left\{1 + \left[\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})\right]/4\right\}^2 - \alpha_{00}(\theta_{00}) = 0 \quad (20)$$
$$\eta_0(\theta_0)\left\{1 + \left[\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})\right]/4\right\}^2 - \eta_{00}(\theta_{00}) = 0 \quad (21)$$

First, notice that if $\theta_0 = 0$ then $\alpha_0(\theta_0) = 0$ and θ_{00} must be positive for (20) to be verified. Similarly, if $\theta_{00} = 0$ then $\alpha_{00}(\theta_{00}) = 0$ and θ_0 must be positive for (20) to be verified. Second, notice that if $\theta_0 \to \infty$ then $\eta_0(\theta_0) \to 0$ and hence it must be that $\theta_{00} \to \infty$ for (21) to be satisfied. Similarly, if $\theta_{00} = 0$ then $\eta_{00}(\theta_{00}) \to \infty$ and then it must be that $\theta_0 = 0$ as well.

By using the implicit function theorem, totally differentiate (20) to get

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

where $A = 1 + [\alpha_1(\theta_{00}) - \alpha_{00}(\theta_{00})]/4$ and $1 < A < 5/4$ since $\alpha_1(\theta_{00}) > \alpha_{00}(\theta_{00})$. Rewriting

$$\frac{d\theta_0}{d\theta_{00}} = -\frac{A\alpha_0(\theta_0)\alpha_1'(\theta_1(\theta_{00}))/2 + \alpha_{00}'(\theta_{00})[1 - A\alpha_0(\theta_0)/2]}{\alpha_0'(\theta_0)A^2} \tag{22}$$

Both the denominator and the numerator in (22) are positive, thus $d\theta_0/d\theta_1 < 0$ for all θ_1. Then rewrite (21) as follows

$$\frac{\alpha_0(\theta_0)}{\theta_0}\{1 + [\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})]/4\}^2 - \frac{\alpha_{00}(\theta_{00})}{\theta_{00}} = 0$$

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

Totally differentiate (23) to get

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

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

To finish. Since $\eta_1 - \eta_{00}x = 0$, it follows that $\frac{\alpha_1}{\theta_1} - \frac{\alpha_{00}x}{\theta_{00}} = 0$

Using (20), $4k/x - \alpha_0(\theta_0)\{1 + [\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})]/4\}^2 - \alpha_{00}(\theta_{00}) = 0$

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

then $\theta_{00}4k/x - \theta_{00}\alpha_{00}(\theta_{00}) = \theta_{00}\alpha_0(\theta_0)\{1 + [\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})]/4\}^2$.

$$\alpha_0(\theta_0)\theta_{00}\{1 + [\alpha_1(\theta_1(\theta_{00})) - \alpha_{00}(\theta_{00})]/4\}^2 - \alpha_{00}(\theta_{00})\theta_0 = 0 \tag{23}$$

$$\theta_{00}4k/x - \theta_{00}\alpha_{00}(\theta_{00}) - \alpha_{00}(\theta_{00})\theta_0 = 0 \tag{24}$$

$d\theta_0/d\theta_1 < 0$.[8] ■

[8] check relation α' and η for calibration.