What's in an Airbnb Five-Star Rating? An Empirical Model of Bayesian Persuasion*

Christian Decker†

November 7, 2022

Job Market Paper

Click here for the latest version.

Abstract

This paper studies the welfare effects of Airbnb's customer rating system using a structural empirical model of Bayesian persuasion with moral hazard. In 2019, over 71% of Airbnb listings in the United States displayed the highest possible rating of 5 stars. The Bayesian persuasion approach reveals that pooling all 'adequate' qualities above a certain threshold in this 5-star rating expands the set of listings customers may choose over the outside options, thereby increasing Airbnb's market shares and profits. I embed the Bayesian persuasion rating system design problem in a numerically solvable demand model of the short-term accommodation market. Moreover, the model incorporates Airbnb's pricing and the hosts' decision to join the platform and exert costly effort to improve their quality. I exploit variation in the rating distribution and market conditions across 56 major travel destinations in the United States over 2018 and 2019 to structurally estimate this model and back out the distribution of unobserved quality. Counterfactual exercises suggest that Airbnb's strategic rating system design led to a consumer welfare loss of US$288M and a redistribution of profits from high- to medium-quality hosts of almost US$750M compared to fully revealing ratings in the markets and period studied.

*Funding through the UBS Center Scholarship and the Zurich GSE Director's Grant is gratefully acknowledged. I would like to thank my advisors Greg Crawford, Marco Ottaviani, Armin Schmutzler, and Holger Sieg for their invaluable support and guidance. Moreover, I thank Heski Bar-Isaac, Lorenzo Casaburi, Matteo Courthoud, Piotr Dworczak, Mike Egesdal, Alessandro Gavazza, Vardges Levonyan, Inessa Liskovich, Alessandro Lizzeri, Nicola Pavanini, Marek Pycia, Regina Seibel, Julian Teichgräber, and Joachim Voth for helpful discussions and comments. All errors are my own.

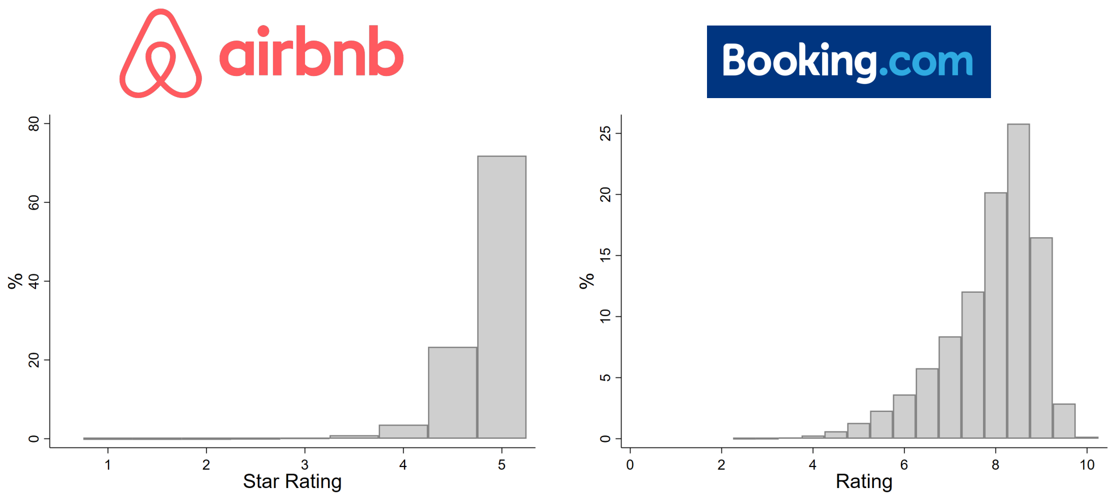
†University of Zurich. E-mail: christian.decker@econ.uzh.ch.

1. Introduction

Two-sided sharing economy online marketplaces, such as Airbnb and Uber, have proliferated over the past years at an impressive rate. Their business model is simple: they match two private persons – one who offers a specific good or service and one who demands it. The backbone of this innovative idea is trust. The platform needs to find a way to create the mutual confidence between two strangers required to engage in such a transaction. For instance, customers might be concerned about their safety when lodging in a stranger's house or riding in a stranger's car. For this purpose, most platforms have established rating systems so customers can review their experience after a transaction. In this way, future platform users can distinguish good and trustworthy participants from bad ones, and everyone is incentivized to exert effort to maintain a good reputation. Hence, one may think that it should be in the platforms' interest that these ratings are as honest and accurate as possible.

However, the rating distributions we actually observe for major sharing economy platforms are far from being fully informative, but in general implausibly positive and highly skewed to the left (see, e.g. Hu, Zhang, and Pavlou, 2009; Zervas, Proserpio, and Byers, 2021). The left panel of Figure 1 shows the distribution of average star ratings of the universe of Airbnb ratings in the US, with an average rating of 4.83 and over 71% of listings receiving the maximum score of 5 stars. Essentially, the

Figure 1: Distribution of Ratings: Airbnb vs Booking



Notes: Left panel: average star ratings of the universe of 714,393 Airbnb listings (with three or more ratings) in the US in December 2019; right panel: average ratings of 35,344 hotels (with three or more ratings) on booking.com in the US in January 2022. Source: Transparent Intelligence and booking.com.

9-step scale from 1 to 5 almost degenerates to a binary signal of 'acceptable' (5 stars) or 'not good enough' (4.5 stars and below).[1] It is remarkable that the distribution of hotel ratings on booking.com in the US, shown in the right panel, does not feature this pattern but is more dispersed and informative.[2] This comparison might appear puzzling, as one would expect that informative ratings are much more important for the sharing economy platform, where there are no official standards or other sources of information.

This paper quantifies consumer, host, and platform welfare effects of Airbnb's rating system using a structural empirical model of Bayesian persuasion with moral hazard (Kamenica and Gentzkow, 2011; Boleslavsky and Kim, 2021).[3] The Bayesian persuasion framework is uniquely well suited to study these effects as it highlights the role of rating systems in impacting customers' perception of the platform relative to other alternatives (here hotels or staying at home) on top of its role in incentivizing costly host effort to maintain good quality, which is often the focus of the literature on rating systems. This approach reveals that the intuitive reasoning in the opening paragraph is incomplete. While the rating system is crucial to filter out the lowest qualities, the uninformative distribution of ratings for listings above a certain quality threshold, which we observe in the data, may actually be beneficial for profit-maximizing platforms. By pooling all 'adequate' qualities above this threshold in the highest possible rating of 5 stars, such a design expands the set of listings customers may choose over the outside options, thereby increasing Airbnb's market shares and profits. This *upper-censoring* of information on quality harms customers on average, even though they rationally update their beliefs according to Bayes' rule knowing perfectly which qualities are included in the 5-star rating.

There are two primary contributions of this paper. First, the Bayesian persuasion approach sheds light on Airbnb's (and, more generally, sharing economy platforms') incentives for strategic information disclosure through the design of rating systems and the interaction of these incentives with the platform's pricing decision and hosts' decision to join the platform and exert effort to improve their quality. My structural estimation of a model that incorporates all these aspects allows me to quantify

[1] The distribution of star ratings for Uber drivers reported by Athey, Castillo, and Chandar (2021) displays the same pattern with over 80% of 5-star ratings.

[2] Between December 2019 and January 2022, there have not happened any modifications to either of the two rating systems that would explain the different shapes of the distributions.

[3] For related foundational literature on information design under sender commitment, see also Aumann, Maschler, and Stearns (1995), Calzolari and Pavan (2006), Ostrovsky and Schwarz (2010), and Rayo and Segal (2010).

the welfare impact of Airbnb's rating system design for the platform, hosts, and consumers.

More broadly, this paper also speaks to the ability of digital platforms to extract and control large amounts of data and information. In the current market environment, ratings on their own platform are customers' only relevant source of information about the unobserved quality of a specific listing. This gives platforms the power to use ratings as a tool for persuasion to their benefit and the harm of customers and certain groups of sellers. In settings where there are alternative information sources, such as for hotels,[4] booking platforms do not have this power, and we observe much more dispersed and informative rating distributions like the one in the right panel of Figure 1.[5] As such, my welfare comparison of the current Airbnb rating system design relative to a full information counterfactual provides a first indication of the benefits regulatory interventions that break this 'information monopoly' could achieve.[6] An example of such an intervention could be the introduction of an independent platform collecting ratings of Airbnb listings.

The second contribution of this paper is methodological. To my knowledge, this work is the first structural estimation of a Bayesian persuasion model. I embed the Bayesian persuasion information design problem in a discrete choice demand model and combine it with other supply-side features (host effort to maintain quality, host participation, and platform pricing) to match the key features of the short-term accommodation market. The critical challenge in this regard is to specify a model that is rich enough to adequately reflect market interaction but is still tractable. I develop a novel numerical solution algorithm, based on the duality approach to Bayesian persuasion (Dworczak and Martini, 2019; Dworczak and Kolotilin, 2019), that allows me to solve this model at a low computational cost and estimate its primitives via the generalized method of moments exploiting cross-market variation in exogenous conditions and predicted outcomes. This toolkit may also be used to analyze rating systems of other platforms, such as Uber, or study empirically other applications of information design that feature similar censoring information structures, such as grade inflation in schools (Ostrovsky and Schwarz, 2010; Boleslavsky and Cotton,

[4]Such alternative sources for information on specific hotels include competing booking platforms (booking.com, Expedia, etc.), but also ratings on third-party platforms such as TripAdvisor and Google.

[5]This reasoning is reminiscent of Lizzeri (1999) who shows that a monopoly certifying intermediary may have incentives to only reveal whether the quality is above some minimal standard, while competition among certifiers can lead to full information revelation.

[6]Recently, customer ratings in online markets have caught competition authorities' attention, for instance, in Germany (Bundeskartellamt, 2020).

2015).

At the core of my structural model is the platform's information design problem of choosing a profit-maximizing information structure. In this context, this means that Airbnb determines the mapping from unobserved quality to star ratings. As discussed in more detail in Section 3.1, there are several design elements the platform can use to influence this mapping.

The objective function of the Bayesian persuasion information design problem derives from a discrete choice demand system that features a unimodal distribution of reservation values for choosing Airbnb over hotels and the outside option of staying at home. In this unimodal case, an *upper-censoring* information structure is optimal for the platform.[7] That is, quality is fully revealed below a certain threshold (empirically, gradual steps from 1–4.5 stars), and above this threshold, all qualities are pooled into one single signal–empirically, the 5-star rating. In Section 2, I discuss in more detail the intuition behind this result and its welfare consequences, and I show that comparative statics with respect to important market conditions like prices and substitution elasticities result in different quality-cutoffs for the pooling signal and a different share of 5-star ratings. In Section 3.3, I will discuss how these comparative statics can help explain reduced-form patterns observed in the data and how cross-market variation can be used to infer the shape of the unobserved quality distribution.

I embed this Bayesian persuasion problem of designing the profit-maximizing rating system in an empirical model of the short-term accommodation market (Section 4). The platform chooses the information structure and prices, a representative host decides which properties to list on Airbnb and how much costly effort to exert to maintain quality, and customers choose between Airbnb, hotels, and the outside option of staying at home. Customers' discrete choice problem is strongly interconnected with the platform's information design problem. The posterior expected Airbnb quality, depending on the information structure, enters the expected indirect utility for Airbnb. Vice-versa, the conditional choice probability function for Airbnb, which enters the objective of the Bayesian persuasion problem, is derived from the demand system.

To solve this model, equilibrium conditions can be restated as additional constraints of the platform's information design problem. Building on very recent insights from

[7] As explained in detail in Section 2, this upper-censoring solution for Bayesian persuasion problems with S-shaped posterior means objective function has first been derived by Kolotilin (2018).

economic theory, particularly from the duality approach to Bayesian persuasion (Dworczak and Martini, 2019; Dworczak and Kolotilin, 2019), I develop a novel numerical solution algorithm to solve this complex problem efficiently in milliseconds. Being able to solve this model repeatedly at a low computational cost allows me to estimate the whole empirical model jointly by the generalized method of moments with data from 56 major travel destinations in the US for the years 2018 and 2019, under the assumption that observed market outcomes (e.g., prices, the share of 5-star ratings, market shares) correspond to the optimal solution of this Bayesian persuasion problem in each market and year.

Counterfactual exercises reveal that the platform would not be able to function without any rating system that identifies the worst qualities and allows it to establish a minimum acceptable quality and safety standard. This mechanism works well with the platform-optimal upper-censoring design (which, by my identifying assumption corresponds to the currently implemented design) featuring informative ratings at the bottom of the distribution.

Compared to a counterfactual scenario with fully revealing ratings, the platform-optimal upper-censoring ratings reduce the average value of Airbnb to customers by 3.5% per booked night or aggregated over the years 2018 and 2019 and the 56 studied geographical areas (covering about 56% of Airbnb listings in the US), US$288.2M. About 11% of this consumer surplus loss can be attributed to upper-censoring ratings providing weaker incentives for hosts to exert costly effort to improve quality. The remaining 89% can be attributed to customers not being able to differentiate 'mediocre' from 'outstanding' Airbnb listings ex-ante, occasionally leading to ex-post 'wrong' choices. Moreover, the platform-optimal design increases Airbnb's profits by US$17.9M and aggregate host profits by US$103.0M. Furthermore, it leads to a redistribution of US$747M from high-quality hosts, whose booking probability decreases by being included in the pooling signal, to medium-quality hosts, whose booking quality is increased by the pooling.

The remainder of this paper is organized as follows. Section 1.1 provides an overview of the related literature. In Section 2, I introduce a simple stylized persuasion model that illustrates the basic idea of why a profit-maximizing rating system does not feature perfect information transmission and how this affects hosts' and customers' welfare. In Section 3, I elaborate on the empirical context and data, I argue how the design of rating systems implies the choice of an information structure in the sense of a Bayesian persuasion model, and I present reduced-form cross-market patterns. Section 4 presents the structural model. Section 5 elaborates on the estimation method and identification, and Section 6 presents the results of the structural estimation. Section 7 considers counterfactual experiments with fully revealing ratings and without any ratings to quantify the welfare impact of strategic rating system design. Section 8 concludes.

1.1. Related Literature

This research project brings together several strands of the theoretical literature on information design and the empirical literature on customer ratings. On the one hand, following Kamenica and Gentzkow (2011) in the last decade, the theory of Bayesian persuasion, and more generally, information design, has been studied extensively. Even though this framework has been used to analyze numerous applications theoretically,[8] to my knowledge there is no empirical work building directly on a Bayesian persuasion model.[9] This lack of empirical work might be partially because only very recent theoretical contributions provide more readily applicable conditions for the characterization of the optimal information structure and solution methods, which are applicable in empirically relevant settings with continuous state space and continuous action space of the receiver (Kolotilin, 2018; Dworczak and Martini, 2019; Kolotilin, Mylovanov, and Zapechelnyuk, 2022) and additional constraints (Dworczak and Kolotilin, 2019; Boleslavsky and Kim, 2021; Doval and Skreta, 2022). These results from the current frontier of theoretical research constitute the basis for my empirical model and the novel numerical solution algorithm I develop.

The previous theoretical literature on reputation in online markets focuses mainly on the role of rating systems in maintaining quality and preventing moral hazard and adverse selection; see, for instance, Tadelis (2016) and Klein, Lambertz, and Stahl (2016).[10] However, this strand of literature does not consider that platforms may exploit a strategic design to influence customers' willingness to choose the platform over competitive alternatives.

[8] See Kamenica (2019), Bergemann and Morris (2019), and Bergemann and Ottaviani (2021) for extensive overviews of the current state of the literature. The applied theory papers closest related to mine are the applications to matching markets (Romanyuk and Smolin, 2019), grading in schools (Ostrovsky and Schwarz, 2010; Boleslavsky and Cotton, 2015), financial markets (Duffie, Dworczak, and Zhu, 2017), and media censorship (Kolotilin, Mylovanov, and Zapechelnyuk, 2022).

[9] Dranove and Jin (2010) provide a survey of earlier theoretical and empirical papers on certification and quality disclosure before the recent advances in the theory of information design in the last decade.

[10] Relatedly, Rossi (2021) investigates theoretically and empirically how competition between Airbnb hosts affects the role of ratings in encouraging host effort.

The most closely related approach to this paper in studying the design of a rating system as the choice of an information structure is taken in a series of recent theoretical contributions by Saeedi and coauthors. Hopenhayn and Saeedi (2022b) derive welfare-optimal rating systems in competitive markets with adverse selection based on vertical seller quality but with an exogenous prior distribution. Hopenhayn and Saeedi (2022a) solve the same problem when the planner can use only a limited number of signals. Saeedi and Shourideh (2020) study the design of rating systems in the presence of both adverse selection and moral hazard. Similar to this paper, their model results in an information design problem with endogenous prior. However, their analysis differs from my model in that prices and costs for quality provision are increasing in the vertical quality level, and the maximization objective is a weighted average of the surplus of different types of sellers. Vellodi (2021) studies the role of ratings as a barrier to entry and how particular designs can alleviate this problem.

On the other hand, on the empirical side of information economics, Vatter (2021) studies the design of US Medicare Advantage provider quality scores and eventually solves for the rating design that optimizes market outcomes. In contrast to this paper, he estimates supply and demand responses to the rating design from exogenous changes in how rating scores are calculated and does not assume that the observed outcomes result from an optimal or profit-maximizing information disclosure problem. To my knowledge, the most closely related work on structural estimation of information disclosure games, which relies on an optimality assumption of observed disclosure patterns, is found in the accounting and finance literature, where variants of the Dye (1985) model are estimated (Bertomeu, Ma, and Marinovic, 2020; Bertomeu, Marinovic, Terry, and Varas, 2022). Fréchette, Lizzeri, and Perego (2022) study in lab experiments if different communication games, including Bayesian persuasion, lead to the theoretically expected outcomes regarding information transmission.

Moreover, there is empirical work on the determinants of online rating informativeness and rating inflation. This literature is mainly concerned with the specific mechanisms which can lead to overly positive ratings (Luca, 2016; Schoenmüller, Netzer, and Stahl, 2018; Filippas, Horton, and Golden, 2022; Fradkin, Grewal, and Holtz, 2021).[11] However, these studies do not address platforms' incentives to explicitly discourage more informative ratings.

The most closely related empirical papers on the competition between Airbnb and hotels are Farronato and Fradkin (2022) and Schaefer and Tran (2020), but their

[11] Relatedly, Mayzlin, Dover, and Chevalier (2014) and He, Hollenbeck, and Proserpio (2021) study outright manipulation of ratings, which can also lead to inflation.

demand estimations and welfare analyses do not consider the role of ratings and the design of rating systems as a choice variable of the platform.

2. The Basic Mechanisms

In this section, I introduce a simple stylized Bayesian persuasion model that allows me to illustrate the basic intuition of why a profit-maximizing rating system does not feature perfect information transmission and how this affects hosts' and customers' welfare. For this purpose, I abstract from ratings' role in disciplining hosts' moral hazard and focus exclusively on their role in persuading future customers. Moreover, I keep constant host participation, prices, and observable quality characteristics, which will all be moving parts in my structural model presented in Section 4.

On top of observable quality characteristics like location, size, and amenities, each Airbnb listing has an unobserved quality component $\omega \in \mathbb{R}$ that adds to the utility a customer derives from booking this listing. Within a market, this unobserved quality component is distributed according to distribution F with density $f(\omega)$. Throughout the paper, I focus on the case where F is a normal distribution.[12] While potential customers know the prior distribution F, the only informative signal they have about the unobserved quality component of a specific listing is its rating $s \in \mathcal{S}$, where the signal space $\mathcal{S}$ can be interpreted as the set of possible ratings (e.g., 1-5 stars).

The fundamental assumption to model Airbnb's design of the rating system as a Bayesian persuasion problem is that the platform can choose and commit to the information structure. In this context, this means that Airbnb can fully determine the mapping from unobserved quality ω to ratings s. As discussed in more detail in Section 3.1, this assumption can be justified by several design elements the platform can use to influence this mapping. Formally, an information structure is a measurable mapping $\pi : \mathbb{R} \to \Delta(\mathcal{S})$.

There is a unit mass of risk-neutral potential customers who consider booking Airbnb listings. Given their prior F and information structure π, for every rating $s \in \mathcal{S}$, they form their posterior belief about the unobserved quality ω of a listing according to Bayes' rule. The posterior expected quality enters additively customers' expected utility for booking an Airbnb listing. Customers will choose Airbnb if their expectation of the unexpected quality component lies above their reservation value

[12] All the results in this section do not depend on the normality assumption, but hold true with any prior distribution with a strictly positive density (Kolotilin, Mylovanov, and Zapechelnyuk, 2022). The normality assumption will be needed to keep the structural model with endogenous host effort tractable (Section 4).

$r \in \mathbb{R}$. These reservation values of potential customers are distributed according to distribution H with density $h(r)$. In my structural model, this distribution of reservation values will be derived from a discrete choice demand model in which customers may choose between Airbnb, hotels, and staying at home. Throughout the paper, I focus on the case in which the distribution of reservation values is unimodal. A unimodal distribution with more customers with reservation values in an intermediate range than at the extremes appears natural in most applications. Commonly studied demand systems in modern Empirical Industrial Organization (logit, nested logit, and in many cases also mixed logit) yield such a unimodal distribution as well.

If ratings were perfectly informative (i.e., if each quality was mapped into a different signal), the probability that a customer chooses Airbnb, which equivalently can be interpreted as the expected market share of Airbnb, would be given by

$$\begin{aligned} \rho^A_{full} &= \int_{\mathbb{R}} f(\omega) \left[\int_{-\infty}^{\omega} h(r)dr \right] d\omega \\ &= \int_{\mathbb{R}} \underbrace{H(\omega)}_{=\rho^A(\omega)} dF(\omega). \end{aligned}$$

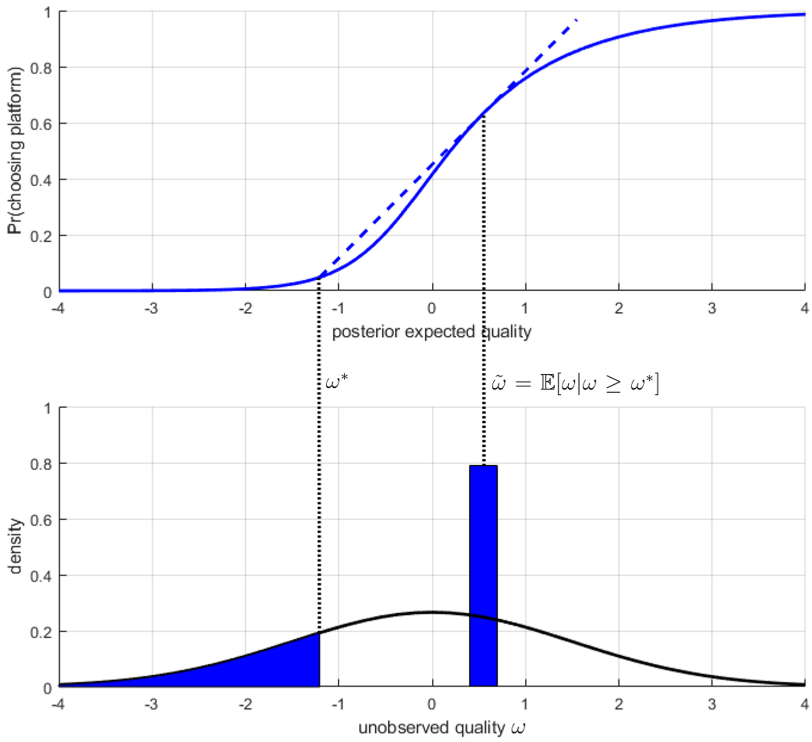
As in this discrete choice context the cumulative distribution function of reservation values H can also be interpreted as the conditional choice probability function for Airbnb depending on posterior means m, in the following, I will refer to it as $\rho^A(m)$.

The Bayesian persuasion approach reveals that if Airbnb wants to maximize its expected market share ρ^A (which for fixed prices is equivalent to maximizing profits), it can do better than fully revealing ratings. As the expected market share of an information structure π depends on posterior beliefs only through the distribution of posterior means it induces, it is convenient to treat the distribution of posterior means G directly as the choice variable of the platform. Therefore, the platform's Bayesian persuasion problem of designing a profit-maximizing rating system can be written as

$$\max_G \int_{\mathbb{R}} \rho^A(m)dG(m)$$

subject to the constraint that the prior F is a mean-preserving spread of the distribution of posterior means G, which captures that posteriors must be consistent with Bayesian updating (Blackwell, 1953; Kolotilin, 2018; Gentzkow and Kamenica, 2016).

Figure 2: Upper-Censoring Solution for S-Shaped Objective Function



Notes: The solid blue line in the top panel plots the probability that a customer chooses Airbnb over other options as a function of the posterior mean of believed unobserved Airbnb quality, $\rho^A(m)$, derived from a nested logit demand system. The dashed blue line corresponds to the optimal price function of the dual problem (Dworczak and Martini, 2019). The bottom panel plots the prior density of unobserved quality $f(\omega)$ (black line) and the distribution of posterior beliefs under platform-optimal upper-censoring ratings (blue mass). Under platform-optimal upper-censoring ratings, unobserved Airbnb quality ω is fully revealed below the threshold ω^* (empirically, gradual steps from 1–4.5 stars), and above this threshold, all qualities are pooled into one single signal–empirically, the 5-star rating–that induces the posterior expected quality of the conditional expectation of the distribution above this threshold $\tilde{\omega} = \mathbb{E}[\omega|\omega \geq \omega^*]$.

Figure 2 illustrates this Bayesian persuasion problem graphically. The black line in the bottom panel depicts the density $f(\omega)$ of unobserved Airbnb quality. The solid blue line in the top panel plots the conditional choice probability for Airbnb as a function of posterior expected quality, $\rho^A(m)$. Recall that this conditional choice probability function coincides with the cumulative distribution function of reservation values for the choice of Airbnb. As the distribution of reservation values is unimodal, this conditional choice probability function is *S-shaped*.[13]

In the case of an S-shaped conditional choice probability function, an *upper-*

[13] A function $h(x)$ is *S-shaped*, if there exists a unique inflection point $\bar{x}$ such that h is convex for all $x < \bar{x}$ and concave for all $x \geq \bar{x}$.

censoring information structure makimizes the platform's market share and profits. That is, quality is fully revealed below a certain threshold ω^* (empirically, gradual steps from 1–4.5 stars), and above this threshold, all qualities are pooled into one single signal (empirically, the 5-star rating) that induces the posterior expected quality of the conditional expectation of the distribution above this threshold $\mathbb{E}[\omega|\omega \geq \omega^*]$.

Intuitively, for the initial convex portion of the conditional choice probability function (with a decreasing density of reservation values), pooling would harm higher qualities more than lower qualities would benefit, leading to full revelation being superior for this part of the distribution. For the concave portion of the choice probability function (with an increasing density of reservation values), however, pooling is superior since lower qualities benefit more than higher qualities are harmed. Taken together, for the convex-concave S-shape, this results in the described upper-censoring structure where the cutoff is determined by the trade-off of improving the choice probability of lower qualities by including them into the pooling signal as long as it does not dampen the choice probability of the infra-marginal higher qualities too much.

Formally, the cutoff ω^* is implicitly defined by

$$\frac{\rho^A(\tilde{\omega}) - \rho^A(\omega^*)}{\tilde{\omega} - \omega^*} = (\rho^A)'(\tilde{\omega}) \tag{1}$$

with $\tilde{\omega} = \mathbb{E}[\omega|\omega \geq \omega^*]$ and $(\rho^A)'(m) = \frac{d\rho^A(m)}{dm}$. This result for S-shaped posterior mean objective functions was first shown by Kolotilin (2018), but the duality approach proposed by Dworczak and Martini (2019) also allows a straightforward derivation of this solution. The derivation of my novel solution algorithm for my structural model (Appendix B.2) builds on this latter approach because it generalizes to persuasion problems with additional constraints (Dworczak and Kolotilin, 2019).

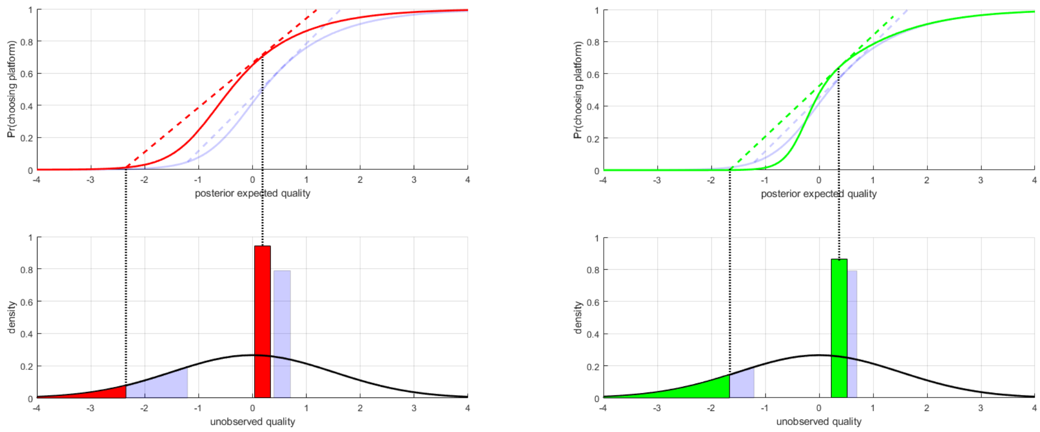
Kolotilin, Mylovanov, and Zapechelnyuk (2022) solve the same persuasion problem with an S-shaped objective function with a first-order approach which allows for a nice intuitive explanation of why upper-censoring is indeed optimal. In the setting of Figure 2, consider an alternative objective function that coincides with the actual S-shaped objective up to ω^*, but above ω^* is given by the dashed straight line. This alternative objective function is convex and therefore allows for an interpretation as the Bernoulli utility function of a risk-loving agent. Hence, under this alternative objective, fully revealing signals, which induce the riskiest lottery over m, are optimal. The construction of the dashed line according to equation 1 with tangency to the S-shaped function at $\tilde{\omega}$ guarantees that the value of the alternative objective under fully revealing signals coincides with the value of the actual objective under the suggested upper-censoring signals. As the alternative objective is weakly greater than the actual objective at any point, it is not possible that the actual objective can assume a value higher than the maximum of the alternative objective. Hence, the upper-censoring information structure must be optimal.

Welfare Consequences. This stylized model also allows for an intuitive illustration of the welfare consequences of upper-censoring ratings for hosts of different quality and customers relative to full information, which I will be able to quantify with my structural model. Hosts with quality above the expectation of the pooling signal $\tilde{\omega}$ are actually harmed by the upper-censoring as their probability of being booked decreases. Consequently, also the revenue Airbnb can generate from these hosts decreases. However, this loss is more than made up for by the increased booking probability for listings with intermediate quality $\omega \in [\omega^*, \tilde{\omega}]$. This is the case because, with a unimodal distribution of reservation values, there are more customers with reservation values in this intermediate range, whom the upper-censoring newly persuades to book on the platform, than there are customers with very high reservation values, whom the platform loses by the pooling of qualities. Platform profit and average host profits are higher under upper-censoring, but the latter average effect hides a substantial redistribution of profits from high-quality to medium-quality hosts.

Since customers enter a 'lottery' when booking a 5-star listing, some will end up with a realized quality above their expectation and some below. However, these differences cancel each other out exactly when aggregating over all customers due to the Bayesian updating constraint. Instead, actual welfare losses derive from a distortion of choices relative to the full information scenario. Consider a customer with reservation value slightly below $\tilde{\omega}$, say 0, in the setting of Figure 2. Such a customer chooses Airbnb when seeing the pooling signal, but he might end up with a lower realized quality, for instance, -1, and if he had known this quality ex-ante, he would not have chosen Airbnb but one of the other options, and he would have been better off. Vice-versa, consider a customer with reservation value above $\tilde{\omega}$, say 3. Such a customer will never choose Airbnb because a 5-star rating is not enough to convince him. However, if he could have identified the very best quality listings, he may have wanted to choose Airbnb, and he would have been better off than with his actual choice.

Comparative Statics. My structural estimation will rely on the assumption that

Figure 3: Comparative Statics of the Upper-Censoring Solution



Notes: Left panel: comparative statics with respect to the ex-ante attractiveness of the platform; right panel: comparative statics with respect to the degree of substitutability between the platform and other options. In light blue in the background of both panels, the baseline scenario from Figure 2 is repeated. The conditional choice probability functions in these examples are derived from a nested logit demand system with Airbnb and hotels in one nest, similar to the specification introduced in Section 4. In the left panel, the Airbnb price is reduced. In the right panel, the nesting parameter, which captures the correlation in idiosyncratic tastes for Airbnb and hotels, is increased.

Airbnb chooses the optimal information structure in different markets in response to the respective market environment. Therefore, it is vital to understand how different market conditions influence the solution of the platform's information design problem. For this purpose, I will now build on the stylized model with an S-shaped objective function from Figure 2 and discuss intuitively some basic comparative statics for the upper-censoring solution.[14] The two graphs in Figure 3 visualize how the cutoff for upper-censoring, and thus the share of 5-star ratings, change with differently shaped conditional choice probability functions resulting from different market conditions. These comparative statics with respect to the ex-ante attractiveness of the platform and the degree of substitutability with other options give a first idea of how cross-market variation can be used to infer the shape of the unobserved quality distribution.

The left panel considers the case in which the platform is relatively more attractive ex-ante compared to the baseline scenario from Figure 2. As a point of reference, this baseline scenario is displayed again in light blue in the background. Examples for such an increase in Airbnb's ex-ante attractiveness include lower Airbnb prices relative to hotels or a lower average hotel quality in a market. Customers are more

[14] Kolotilin, Mylovanov, and Zapechelnyuk (2022) discuss comparative statics for the solution of Bayesian persuasion problems with S-shaped posterior mean objective functions more formally.

likely to choose the platform for any given posterior quality. This is represented in a leftward shift of the conditional choice probability function. As the platform is already more attractive *on average*, it needs to disclose less information to persuade customers.[15] The threshold quality moves to the left, more lower qualities are pooled with high qualities, and we have more 5-star ratings.

The right panel considers the case of a higher degree of substitutability between Airbnb and hotels compared to the baseline scenario from Figure 2 in light blue in the background. Reservation values are shifted from the extremes towards an intermediate range, and the conditional choice probability function gets steeper in this intermediate range. In this case, more lower qualities can be included in the pooling signal without harming the higher qualities too much. The share of 5-star ratings increases.

3. Background, Data, and Reduced-Form Patterns

3.1. The Airbnb Rating Systems

Airbnb, founded in 2008, is the biggest peer-to-peer marketplace for short-term accommodation with a revenue of US$ 6.0B in 2021, generated from 300.6M booked nights of 6M listed properties worldwide (Airbnb, 2022). It provides a platform to match supply and demand, process payments, and ensure safe transactions, not least through its rating system that allows previous customers to share their experience and satisfaction with a specific listing.

During my sample period of the years 2018 and 2019, Airbnb displayed average customer ratings of a listing on a five-star scale rounded to the nearest half-star. Effectively, this yields a 9-step scale from 1, 1.5, ..., to 5 stars.[16] In particular, a listing exhibited 5 stars if its average rating is at least 4.75.

I assume that star ratings are the only informative signals through which potential guests can learn about an unobserved quality component of a specific Airbnb listing.[17]

[15] The economic intuition behind this result is very similar to the applications of information design under sender commitment to grading in schools (Ostrovsky and Schwarz, 2010; Boleslavsky and Cotton, 2015). Harvard can apply a very coarse grading scheme as the students are considered good enough *on average* to be qualified for a good job even without much further information through grades. Low-ranked universities, instead, optimally apply a much more granular grading scheme such that at least the best students can be identified as qualified for a good job.

[16] Towards the end of 2019, Airbnb started testing and eventually switched entirely to displaying average ratings with the precision of two decimal places. See https://www.airbnb.com/resources/hosting-homes/a/in-case-you-missed-it-3-product-updates-to-catch-up-on-114 (accessed on September 11, 2021).

[17] I abstract from the role of 'superhost'-badges and textual reviews. I deem star ratings more

This unobserved quality component may include aspects like the property's cleanliness and kindness and availability of the host, but it does not include anything that could be learned from the description of a listing (e.g., location, amenities, size and type of the property, etc.). The distribution of this unobserved listing quality may feature a certain natural variance that goes beyond what yields a direct monetary reward to hosts or is directly incentivized by the platform, for instance, due to differences in hosts' intrinsic pride, motivation, kindness, or greed.

The key assumption for modeling the design of a rating system as a Bayesian persuasion problem is that the platform can freely choose the information structure. In the context of ratings, this means that Airbnb can perfectly determine the mapping from unobserved quality to star ratings. Loosely speaking, Airbnb sets the 'rules' for ratings, and the customers comply.

There are several elements in the design of the rating system that Airbnb can use to explicitly or implicitly encourage accurate or biased reviews and therefore justify this assumption.[18] While the platform might not be able to pin down the aforementioned conditional probabilities precisely with these design elements, it can influence them substantially toward the desired direction.

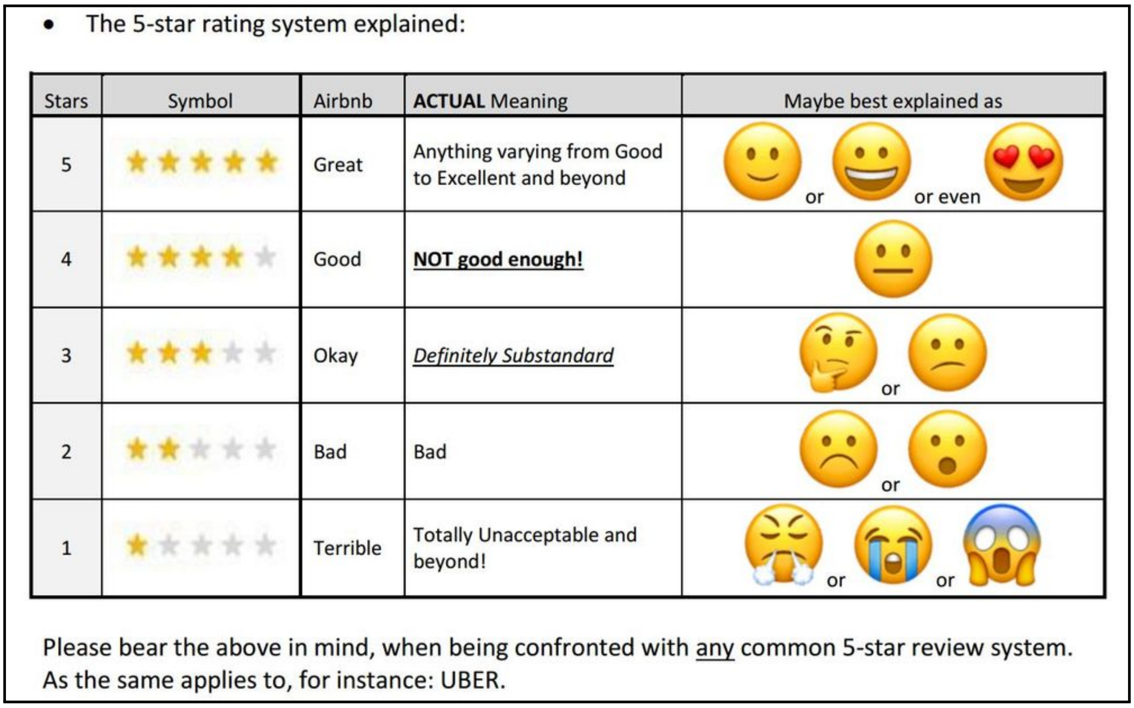
The most powerful design element works by attaching severe negative consequences to bad ratings. Airbnb threatens hosts with exclusion from the platform if their average rating falls below a certain threshold (see Appendix Section D.1 for some anecdotal evidence). These critical thresholds are not announced publicly, but folk wisdom on the internet says that this critical value is around 4.6 and, importantly for my empirical strategy, varies by city.

By attaching such serious consequences to ratings, the platform prescribes a certain interpretation of the ratings, which eventually pins down the information structure. That is, 5 stars mean 'anything varying from good to excellent and beyond', and everything below is different levels of 'not good enough'. These attached consequences induce Airbnb hosts to 'educate' their guests about the meaning of ratings and encourage them heavily to obey the platform's preferred rating scale,

important as they are usually the first anchor point customers look at and are easily comparable across listings. Reading a representative sample of textual reviews of several listings would be much more time-consuming. My model is only concerned with vertical quality, for which I deem the average star rating to be a better proxy, while individual textual reviews might be more informative about the quality of the match with a specific customer; thus, horizontal quality.

[18] See the review paper by Luca (2016) for an overview of other design elements of rating systems that might influence the informativeness of ratings, including the role of reciprocity of ratings, fake/promotional ratings, and the induced social distance/proximity between the two parties of a transaction.

Figure 4: Anecdotal Evidence: Rating Instructions by Airbnb Hosts



- The 5-star rating system explained:

Stars	Symbol	Airbnb	**ACTUAL** Meaning	Maybe best explained as
5	★★★★★	Great	Anything varying from Good to Excellent and beyond	or or even
4	★★★★☆	Good	**NOT good enough!**	
3	★★★☆☆	Okay	*Definitely Substandard*	or
2	★★☆☆☆	Bad	Bad	or
1	★☆☆☆☆	Terrible	Totally Unacceptable and beyond!	or or

Please bear the above in mind, when being confronted with any common 5-star review system.
As the same applies to, for instance: UBER.

Source: https://community.withairbnb.com/t5/Hosting/How-can-we-Hosts-quot-educate-quot-Guests-about-how-the-AirBnB/td-p/680774/page/3 (accessed on September 24, 2022).

for instance, through 'customer instructions' in leaflets they leave for their guests. One example of such a note, found in an online forum for Airbnb hosts, is shown in Figure 4. More are collected in Appendix Section D.1.[19]

Given this influence, it appears reasonable that Airbnb can choose and commit to an information structure in the sense of a Bayesian persuasion model. While the studied effects of rating systems often focus on their role in mitigating adverse selection and moral hazard, this Bayesian persuasion approach reveals that there is a second effect of ratings impacting customers' perception of Airbnb relative to other alternatives, which the platform can exploit to its benefit.

3.2. Data

This project draws mainly from two commercial data sources: *Transparent Intelligence* (TI) for Airbnb data and *STR* for hotel data.

Airbnb. TI provides web-scraped data from short-term rental platforms, including the Airbnb API. This includes all the publicly available information on the Airbnb

[19]Uber drivers display similar 'information sheets' on the back of the front seats of their cars; see Figure D.4.

homepage, such as listings' locations, amenities, availability, and ratings. Moreover, TI uses Machine Learning to predict actual occupancy and transaction prices for every listing.[20] I obtained from TI monthly aggregated supply and performance data for the universe of *all* individual Airbnb listings in the US for the years 2018 and 2019. Since my analysis focuses on the competition between Airbnb and hotels, I restrict the sample to Airbnb listings that can be considered comparable to a hotel room. These are 'entire homes' (i.e., entire houses not shared with the host) with up to two bedrooms and 'private rooms' (i.e., private rooms in a house shared with the host), which make up 71.1% of all listings. The excluded listings are 'entire homes' with more than two bedrooms and rooms shared with the host.

Hotels. STR is a major accommodation industry data provider that tracks hotels' performance worldwide by conducting periodic surveys asking about hotels' room capacity, daily occupation, and generated revenue. These surveys cover nearly 75% of US hotel rooms, including 97% of chain hotels and most significant independent hotels. For hotels not participating in the survey, STR imputes outcomes based on the performance of similar hotels. From STR, I obtained trend reports for 56 markets in the US containing most major cities and holiday travel destinations, covering more than half of all the Airbnb listings in the US in the period studied.[21,22] The data includes daily room availability, occupation, and prices aggregated at the market level, as well as a breakdown of hotel capacity into six quality scales (from economy to luxury). Moreover, daily occupation rates allow me to calculate the share of *compression nights* within a specific time period. These are nights in which hotels hit their capacity constraints and are commonly defined as nights with an occupation above 90%.

Additional Data Sources. The Airbnb and hotel data is complemented with data from various sources.

In January 2022, I web-scraped from the major hotel-booking platform booking.com customer ratings of more than 94% of all hotels in the US, which will serve as a proxy of hotel quality in the demand estimation and whose distribution is plotted in the right-hand panel of Figure 1.

[20] The critical challenge is to distinguish if a night unavailable in a listing's calendar is blocked by the host or actually booked. TI's algorithms are trained with historical data from times when Airbnb had still made this information public and relies on variables such as the common lead time of bookings.

[21] The exact percentage of all Airbnb listings in the United States covered by the 56 geographical markets varies, depending on the month, between 55.6% and 57.8%.

[22] The market selection was partially constrained by the availability from the side of STR. For instance, data for Las Vegas were not available.

Moreover, I retrieve tourist points of interest (POI) from Open Street Map (as of December 2021), whose relative count in different markets will serve as a proxy for the attractiveness of different travel destinations in the demand estimation.

I rely on the land unavailability measures from Lutz and Sand (2019) as a measure of hotel supply constraints, which will serve as an instrument for hotel prices in the demand estimation.

Rental data from the Department of Housing and Urban Development will serve as a proxy for the opportunity cost of offering a property in the short-term instead of the long-term rental market and as an instrument for Airbnb prices in the demand estimation.[23] As a proxy for opportunity costs of providing additional host effort, I rely on average wages from the U.S. Bureau of Labor Statistics Quarterly Census of Employment and Wages.[24] Finally, I use population and household demographics data from the 2010 US census.

Market Definition and Aggregation. For the definition of geographical markets, I stick to the units proposed by STR, which can be considered the industry standard. Table A.1 lists all 56 markets, which I subdivided into urban and non-urban (i.e., mainly regions at the seaside and around national parks) travel destinations.

I aggregate both Airbnb and hotel data for these 56 STR markets at the yearly level for 2018 and 2019, yielding a total of 112 observations. Note that I work with less granular data in the geographical and the temporal dimension than papers focusing exclusively on estimating demand for short-term accommodation (Farronato and Fradkin, 2022; Schaefer and Tran, 2020). This is because the emphasis of this project is on the impact of the design of the rating system. The key modeling assumption presented in Section 3.1 prescribes that Airbnb can manipulate the information structure at the unit of observation, and customers understand this information structure and respond rationally (5 stars in market A have a different interpretation than 5 stars in market B). It is not realistic that both manipulation and accurate interpretation of the information structure happens at a more local level. Neither is it realistic that this happens at a high temporal frequency since the displayed average ratings are cumulative and therefore adjust only slowly. Hence, my estimation relies mainly on variation across larger geographical markets rather than over time within

[23]In particular, I retrieve 'fair market rents' (FMRs) for a one-bedroom apartment, which correspond to the 40th percentile of the rent distribution at the zip code level. To aggregate to STR markets, I take the average of these FMRs over all covered zip codes weighted by the number of available Airbnb rooms.

[24]The STR markets are assigned the annual average weekly wage (all industries) of the MSA, County, or State with the largest geographical overlap.

Table 1: Summary Statistics

		mean	**sd**	**min**	**p50**	**max**
Hotels	Price [US$]	139.12	34.61	85.06	132.44	262.86
	Demand [M]	10.85	8.82	1.11	9.16	37.66
	Supply [M]	15.26	11.61	1.67	12.65	46.51
	Occupancy	0.69	0.06	0.57	0.69	0.87
	Share Compression Nights	0.08	0.08	0.00	0.06	0.48
	Share $\geq$ upper-midscale	0.55	0.09	0.36	0.55	0.83
	No. Hotels	411.8	277.9	53	389.5	1267
	avg booking.com rating	7.92	0.22	7.28	7.93	8.43
Airbnb	Price [US$]	115.37	33.82	66.57	109.48	201.17
	Demand [M]	1.07	1.21	0.13	0.72	7.65
	Supply [M]	2.10	2.23	0.29	1.38	13.56
	Share 5 Star	0.72	0.08	0.50	0.72	0.88
	# Ratings	26.1	11.4	6.7	26.1	59.9
	Entire Home	0.66	0.14	0.40	0.64	0.96
	Bedrooms	1.3	0.1	1.1	1.3	1.7
	Occupancy	0.50	0.07	0.29	0.50	0.65
	Occupancy (non5star/5star)	0.72	0.10	0.41	0.72	0.96
Other	Unavailable Land Share	32.47	18.31	6.18	29.41	74.06
	Population [M]	2.07	1.93	0.07	1.66	9.40
	log tourist poi per 100k inhab.	3.80	0.71	2.50	3.70	5.63
	Daily Rent [US$]	39.33	14.25	21.33	34.65	94.60
	Daily Wage [US$]	161.71	46.18	85.86	149.43	376.14
	Market Size [M]	62.13	57.82	4.59	49.87	282.03
	Area [1,000 km^2]	320.25	846.90	0.88	67.31	4,221.94

Notes: The table shows summary statistics (mean, standard deviation, minimum, median, and maximum) of all variables used in this paper across 112 markets (56 geographical regions x 2 years).

markets.

Market size is defined proportional to the population in each market, where for the two markets with the most rooms offered (hotels and Airbnb combined) per capita, I set market size equal to the total capacity.

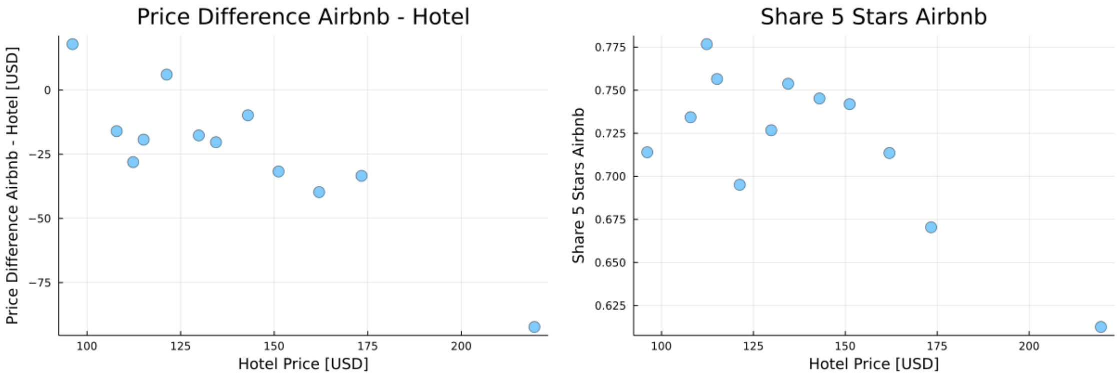
Table 1 shows summary statistics of all relevant variables across the 112 observations. On average, 72% of Airbnb listings have a 5-star rating but across markets this share varies substantially between 50% and 88%. Average Airbnb prices lie about 17% under hotel prices. Hotel capacity exceeds Airbnb capacity by a factor of 7 on average. Losing a 5-star rating decreases a listing's average occupancy by 28%.

3.3. Reduced-Form Patterns

Before moving to the full structural model, in this section, I will show some cross-market correlations of the key variables in my structural model, and I will discuss how the comparative statics of the persuasion model presented in Section 2 and their interplay with other market mechanisms will help to explain some prominent reduced-form cross-market patterns in the data.

Figure 5 shows binned scatter plots of two key endogenous variables of my structural model, the Airbnb-hotel price difference (Airbnb prices minus hotel prices) and the share of Airbnb 5-star ratings, with hotel prices, an impactful shifter of Airbnb's ex-ante attractiveness, on the horizontal axes. The price difference in the left-hand panel of Figure 5 is decreasing almost linearly in the hotel price with a slope less than one in absolute value. Airbnb prices increase on average only by 29 cents for a one-dollar increase in hotel prices.

Figure 5: Cross-Market Correlations of Hotel Prices with *Endogenous* Variables



Notes: The graphs show binned scatter plots of the Airbnb-hotel price difference (left panel) and the share of Airbnb 5-star ratings (right panels) with the average hotel price across 112 markets on the horizontal axes.

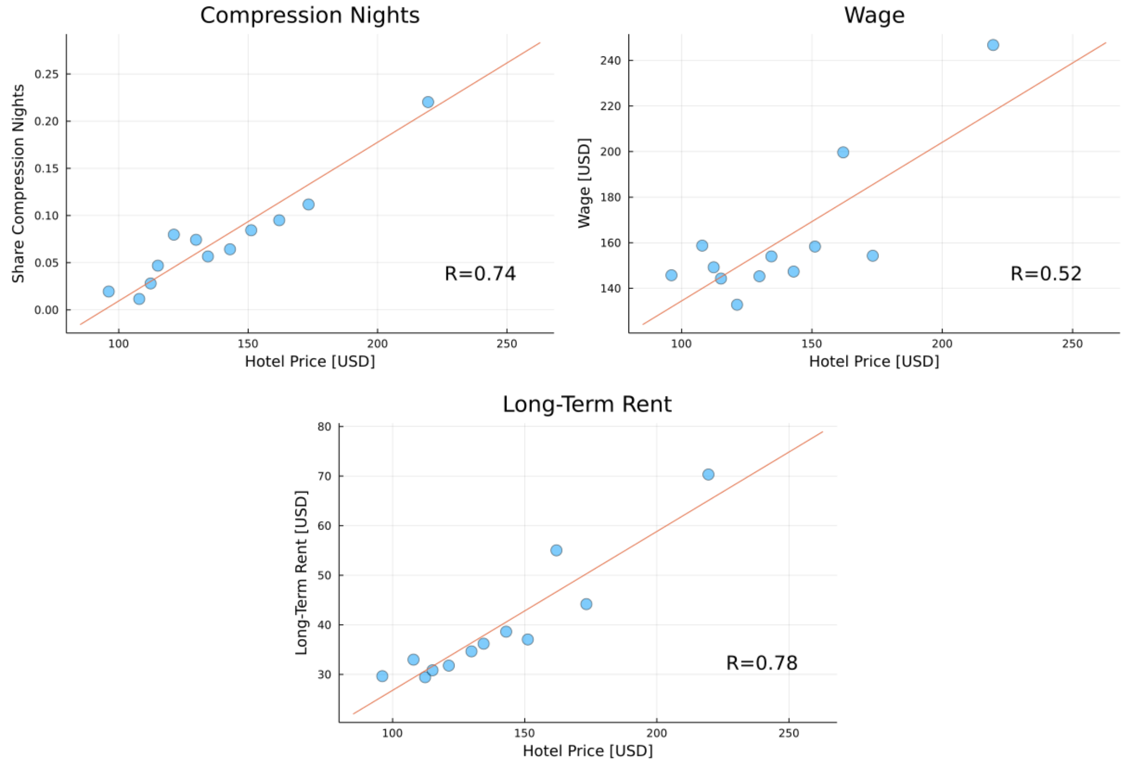
The intuition behind this pattern, which also consistently arises in numerical solutions of my model for a wide range of parameters, is the following. Imagine the hotel price increases. This means that ceteris paribus, Airbnb is more attractive to customers. To increase profits, Airbnb can react in two ways. Either they increase their price as well, or they induce more 5-star ratings as they are more attractive on average and do not need to provide much more additional information to persuade customers (recall the left panel of Figure 3 above). It turns out that the optimal solution is a combination of both. Airbnb increases the price, but by a much smaller margin than the increase in hotel prices; hence, the downward slope in the price difference. In summary, the model features a substitutability between higher Airbnb prices and less information in the form of a higher share of induced 5-star ratings.

Given this substitutability and the discussed comparative statics with respect to the ex-ante attractiveness of the platform, we would expect the share of Airbnb 5-star ratings to be increasing in the hotel price (as higher hotel prices make Airbnb more attractive ex-ante, Airbnb needs to provide less information to persuade consumers). However, empirically (right-hand panel of Figure 5), we observe a hump-shaped relationship between 5-star ratings and hotel prices, with an initial increasing portion for low hotel prices but eventually a sharply decreasing portion for high hotel prices. This pattern arises because the 'attractiveness mechanism' is not the only force at play, but there are additional market mechanisms, which are also reflected in my structural model, pushing in the opposite direction.

To understand the whole picture, we need to consider how hotel prices correlate with other market conditions, which I will treat as exogenous in my model (Figure 6). First, hotel prices have a strong positive cross-market correlation with the share of compression nights per year (top left panel of Figure 6). In compression nights, hotels hit their capacity constraints, and the platform does not face competition from hotels for the marginal customer. Hence, markets with a higher share of compression nights feature a lower degree of substitutability between Airbnb and hotels (when aggregating yearly). Recall from the discussion of the comparative statics in Figure 3 above that a lower degree of substitutability leads to a lower share of 5-star ratings under the profit-maximizing information structure.

Second, hotel prices have a strong positive cross-market correlation with wages (top right panel of Figure 6). Wages may be a shifter of hosts' opportunity costs of devoting more time to increase the quality they can provide, for instance, by cleaning the apartment more thoroughly or interacting more closely with their guests. As such, higher wages may lead to less host effort and a shift of the prior distribution of

Figure 6: Cross-Market Correlations of Hotel Prices with *Exogenous* Variables



Notes: The graphs show binned scatter plots (blue circles), linear fit lines (red lines), and Pearson correlation coefficients ('R') of the share of compression nights (top left panel), daily wages (top right panel), and daily long-term rents (bottom panel) with respect to the average hotel price across 112 markets on the horizontal axes.

qualities to the left. This means that at a higher wage level for a given cutoff quality for upper-censoring, we expect a lower share of 5-star ratings.

Third, hotel prices have a strong positive cross-market correlation with long-term rents (bottom panel of Figure 6). Long-term rents may be a shifter of hosts' opportunity costs of offering their property for short-term rental on Airbnb in the first place. The share of 5-star ratings depends not only on the cutoff for upper-censoring but also (through the denominator) on how many lower-quality listings there are on the platform. On the one hand, with a higher long-term rent level, it may be less attractive to list lower-quality apartments on Airbnb, which are relatively less likely to be booked. On the other hand, with a higher hotel and Airbnb price level, the expected revenue of offering such low-quality apartments, even if booked infrequently, is higher. Taken together, the expected effect of this 'participation mechanism' on the Airbnb 5-star–hotel price relationship is ambiguous.

In summary, we have one potential force pushing towards an increasing cross-market relationship between hotel prices and the share of Airbnb 5-star ratings (the direct 'attractiveness mechanism'), two potential forces pushing towards a negative relationship (the 'substitutability' and 'effort' mechanisms), and one ambiguous force (the 'participation mechanism'). The empirical model I will introduce in the next section is built to allow for all these mechanisms and their non-linear interplay and can replicate the hump-shaped relationship we observe in the data.

4. Structural Model

In this section, I set up a structural empirical model of the short-term accommodation market, focusing on the profit-maximizing design of the Airbnb rating system. Furthermore, I outline a novel numerical solution algorithm for Bayesian persuasion problems with additional constraints like the one arising in this set-up.

Overview. At the core of the model, there is a Bayesian persuasion problem with moral hazard in the spirit of Boleslavsky and Kim (2021). The profit-maximizing platform (sender) designs a rating system (information structure). Unlike in the simple model from Section 2, now the rating system does influence not only the distribution of customers' (receiver) posterior beliefs about quality and hence their decision between booking on Airbnb or another option. In also influences the representative host's (agent) effort choice and hence the actual quality distribution of listings on the platform. The platform's objective function for this Bayesian persuasion problem depends on customers' conditional choice probabilities of the platform relative to a hotel or staying at home, which are derived from a nested logit demand system with some adjustments for hotels' capacity constraints.

Agents. The model features three classes of agents. A continuum of consumers faces a discrete choice problem between booking on Airbnb, a hotel, and the outside option of staying at home or arranging some alternative form of accommodation outside the market (e.g., with friends or relatives). The platform Airbnb designs the rating system by choosing the profit-maximizing information structure and sets profit-maximizing prices. A representative host decides which properties to list on Airbnb and how much costly effort to exert to maintain quality.

Prices. Hotel prices in geographical market n in year t are exogenously fixed at p^H_{nt} and the Airbnb price is $p^A_{nt} = p^H_{nt} + p_{nt}$. Hence, p_{nt} is the price difference between Airbnb and hotels, which may be positive or negative.[25] I assume that the price

[25] I use the price difference relative to hotels instead of price levels as the choice variable for Airbnb as it maps more closely to the comparative statics of the Bayesian persuasion model with respect to the ex-ante attractiveness of Airbnb relative to hotels.

difference p_{nt} is chosen by the platform uniformly for all listings in a market as part of the profit-maximization problem.[26,27]

Airbnb obtains a fixed share $\tau = 0.16$ of the price of every transaction made on the platform as commission.[28] Therefore, for fixed prices, profit maximization implies that the platform aims to maximize the share of customers choosing the platform over the other options.

Quality. The quality of hotels is assumed to be fully observable ex-ante as a function of the STR classification and informative booking.com ratings. The quality of the service provided by the platform depends both on observable characteristics and an unobserved quality component, which, as discussed in the previous section, may include aspects like the cleanliness of the property and the kindness and availability of the host. This unobserved component is modeled as a real-valued random variable ω whose distribution $F_{\underline{\omega}_{nt}}(\omega; e_{nt})$[29] depends on the cutoff quality for hosts' participation $\underline{\omega}_{nt}$ and hosts chosen effort level e_{nt}.

Host Participation and Effort. To capture the aggregate of many small Airbnb hosts' decisions, in each market nt, there is a representative host who owns a portfolio of listings with quality distribution F, where F is normal distribution with mean $\nu + e_{nt}$ and variance σ^2.[30] I assume that the distribution of unobserved quality of

[26] This assumption is backed by Airbnb suggesting prices to hosts and actively encouraging them to comply with these suggestions. See https://community.withairbnb.com/t5/Airbnb-Updates/Airbnb-Answers-Pricing-suggestions/td-p/790645 (accessed on September 11, 2021). Moreover, Airbnb's and hosts' incentives in choosing prices are very much aligned apart from the impact the price has on the hosts' participation decision.

[27] There is a recent literature studying how prices affect ratings of listings through several mechanisms (Carnehl, Stenzel, and Schmidt, 2021; Carnehl, Schaefer, Stenzel, and Tran, 2021), and vice-versa (Teubner, Hawlitschek, and Dann, 2017). Since my analysis is at a higher level of aggregation, I abstract from such dependencies and assume uniform prices within a market. Alternatively, you may interpret p^A_{nt} as a *price index* for market nt that lets actual prices within the market correlate with observable Airbnb quality characteristics. However, to study the rating system design in this Bayesian persuasion framework, prices need to be orthogonal to unobservable listing quality and ratings such that ratings remain the only informative signal for unobserved quality.

[28] The actual Airbnb fee structure in the period studied was more complicated, with service fees split between hosts (about 3%) and guests (about 13%), with the precise amount varying depending on factors like reservation subtotal, the length of stay, and the listing characteristics. As there is, however, no indication that fee structure and levels varied systematically between different regions in the US, I assume, for simplicity, a uniform fee of 16% of the transaction price as paid by the customers. Moreover, any additional costs like cleaning fees should be understood to be part of the Airbnb price p^A_{nt}.

[29] I adopt the convention of identifying probability distributions by their cumulative distribution functions. Moreover, $F_{\underline{\omega}}$ and $f_{\underline{\omega}}$ refer to the cdf and pdf of the distribution resulting from truncating F from below at $\underline{\omega}$.

[30] As I see the variation of the unobserved quality component as a result of differences in hosts' motivation that is drawn from a very large population of potential hosts, I consider the normal distribution a reasonable assumption.

potential Airbnb listings net of the additional host effort component, pinned down by the parameters ν and σ, is the same across all markets.

The representative host faces two decisions. First, he can shift the unobserved quality distribution of his portfolio uniformly to the right by exerting more costly effort e_{nt}. I assume the quadratic cost function $c(e_{nt}) = k_{nt}e_{nt}^2$.

Empirically, I parametrize the marginal effort cost k_{nt} as a function

$$k_{nt} = \kappa_0 + \kappa_1 wage_{nt} \tag{2}$$

of average wage levels in market nt as the monetary opportunity cost of devoting more time to improve the quality of the listing, for instance, by cleaning the property more thoroughly or interacting more closely with the guests.

Second, the representative host decides which properties of his portfolio actually to list on Airbnb. He needs to pay a fixed cost proportional to the mass of listings he decides to offer on the platform. As the expected revenue of a listing is increasing in its quality, the host optimally chooses a cutoff value $\underline{\omega}_{nt}$ such that he offers only qualities above this threshold. The overall fixed cost is given by $\int_{\underline{\omega}_{nt}}^{\infty} T_{nt} f(\omega; e_{nt}) d\omega$ with cost parameter $T_{nt} > 0$.

Empirically, I parametrize T_{nt} as a function

$$T_{nt} = \eta_0 + \eta_1 rent_{nt} + \varepsilon_{nt}^T \tag{3}$$

of daily long-term rent levels in market nt as the monetary opportunity cost of offering a property for short-term instead of long-term rental.[31] The cost shock ε_{nt}^T is unobserved to the econometrician and to the platform.

In summary, the original distribution $F(\omega; 0)$ can be interpreted as the unobserved quality distribution of all *potential* listings in a market without additional host effort incentivized by the rating system. The representative host's participation decision truncates this distribution from the left at $\underline{\omega}_{nt}$, and his effort decision shifts the distribution uniformly to the right by e_{nt}. The resulting truncated and

[31] Farronato and Fradkin (2022), who study demand in the short-term accommodation market over the years 2011 to 2015, suggest market demographics such as the share of single or childless households as shifters for Airbnb supply with the idea that for families it is more costly to host a stranger in their house than for singles. In my more recent sample (years 2018 and 2019), such demographics have very little explanatory power for cross-market variation in Airbnb supply. I attribute this difference to Airbnb shifting more and more from occasional peer-to-peer supply to properties owned by institutional investors for the sole purpose of offering them for short-term rental. Therefore, I work with monetary opportunity costs, such as long-term rent and wages, instead.

shifted distribution $F_{\underline{\omega}_{nt}}(\omega; e_{nt})$ can be interpreted as the empirical distribution of unobserved quality of the *actual* listings on the platform in market nt. While the initial distribution $F(\omega; 0)$ by assumption is the same across all markets, the eventual quality distribution of actual listings $F_{\underline{\omega}_{nt}}(\omega; e_{nt})$ will differ in each market.

Information. Through the design of the rating system, the platform commits to an information structure that determines the ratings of 'previous customers'. 'Previous customers' are not modeled explicitly. Through the design of the rating system, the platform can determine the average rating 'previous customers' give conditional on the quality, which they observe perfectly once having booked a specific listing.[32] Future customers exploit these ratings as the only signal for the unobserved quality ω of a specific offer and update their beliefs according to Bayes' rule. Formally, an information structure is a measurable mapping $\pi_{nt} : \mathbb{R} \to \Delta(\mathcal{S})$ for some signal space $\mathcal{S}$ which can be interpreted as the set of possible ratings (e.g., 1-5 stars). Ex-ante, I do not restrict $\mathcal{S}$.

Customers' Choice. Moving to the demand side of the model, every individual customer i maximizes expected utility by choosing between booking an Airbnb listing (A), a hotel (H), or the outside option of staying at home (0).[33] Customer i's expected utilities for these three options conditional on observing signal $s \in \mathcal{S}$ in market n and year t are given by, respectively,

$$U^A_{int} = \underbrace{\gamma_g + \psi \log(poi_n) - \alpha p^A_{nt} + \boldsymbol{q}^A_{nt}\boldsymbol{\beta}^A + \xi^A_{nt}}_{=\delta^A_{nt}} + \mathbb{E}_{\mu^s_{nt}}[\omega] + \iota_{int} + (1-\zeta)\varepsilon^A_{int}; \tag{4}$$

$$U^H_{int} = \underbrace{\gamma_g + \psi \log(poi_n) - \alpha p^H_{nt} + \boldsymbol{q}^H_{n}\boldsymbol{\beta}^H + \xi^H_{nt}}_{=\delta^H_{nt}} + \iota_{int} + (1-\zeta)\varepsilon^H_{int}; \tag{5}$$

$$U^0_{int} = 0 + \varepsilon^0_{int}.$$

I assume a nested logit structure of the stochastic, idiosyncratic utility shocks with Airbnb and hotels in one nest. That is, $\iota_{int} + (1-\zeta)\varepsilon^A_{int}$, $\iota_{int} + (1-\zeta)\varepsilon^H_{int}$, and ε^0_{int} are assumed to be iid TIEV across the population in each nt and independent of ω. The nesting parameter ζ governs the correlation of preferences between Airbnb and hotels; that is, a larger value of ζ implies a stronger correlation of the idiosyncratic preference components of Airbnb and hotels as opposed to the outside option.

The parameters γ_g represent the value of traveling to a destination in the group g,

[32] I assume that there are sufficiently many 'previous customers' such that there are informative ratings for all listings.

[33] The outside option 0 may also include alternative forms of accommodation organized outside of the market, like lodging a friend's or relative's place.

where g can be urban (u) or non-urban (nu); see Table A.1 for the classification. By poi_n, I denote the count of OSM tourist points of interest (e.g., museums, beaches, memorials, etc.) per 100,000 inhabitants in the region n. The Airbnb and hotel prices in market nt, respectively, are denoted by p^A_{nt} and p^H_{nt}. The vectors $\boldsymbol{q}^A_{nt}$ and $\boldsymbol{q}^H_n$ collect observable Airbnb and hotel quality characteristics. In the case of Airbnb, these are the share of private rooms, entire home studios, and entire home two-bedroom apartments in a market, with entire home one-bedroom apartments as the omitted category. For hotels, these are the average booking.com ratings and the share of hotel rooms in a market classified by STR as *Upper Midscale* and higher.[34]

As standard in the demand estimation literature, ξ^A_{nt} and ξ^H_{nt} are product characteristic or demand shocks, whose realization customers know when making their choice, but which are unobserved to the econometrician. Moreover, I assume that the realizations of these shocks are also unobserved to the platform and the hosts when they make their profit-maximizing decisions. The added-up linear components of Airbnb and hotel utility that are observable to customers are denoted by δ^A_{nt} and δ^H_{nt}, respectively.

The posterior expectation of the unobserved Airbnb quality component ω when observing rating s in market nt is denoted by $\mathbb{E}_{\mu^s_{nt}}[\omega]$.

Hotels' Capacity Constraints. Previous papers estimating demand in the short-term accommodation industry have shown that Airbnb's main impact happens in times and places where hotels are capacity-constrained (Farronato and Fradkin, 2022; Schaefer and Tran, 2020), for instance, in New York City on New Year's Eve. At such so-called *compression nights*, the marginal Airbnb customer, whom persuasive ratings want to target, can effectively only choose between Airbnb and the outside option. This choice between only two options yields different substitution patterns than the choice between all three options and thereby influences the choice of the optimal information structure; recall the intuitive discussion in Section 2. Since the share of compression nights scn_{nt} varies substantially across markets (Figure 6), it is essential to take this into account in the empirical model.

I assume that at compression nights, commonly defined as occupancy above 90%, hotels operate at maximal market share $\bar{\rho}^H_{nt}$ equaling 95% of capacity.[35] On such nights, the marginal customer chooses only between Airbnb and the outside option, yielding an expected Airbnb market share for a given posterior mean $m = \mathbb{E}_{\mu_s}[\omega]$

[34] STR classifies hotel rooms into six classes. These are, from high to low, *Luxury*, *Upper Upscale*, *Upscale*, *Upper Midscale*, *Midscale*, and *Economy*.

[35] As I am looking at a relatively short time horizon, it is reasonable to assume that hotel capacity $\bar{\rho}^H_{nt}$ is exogenous. To keep the demand system tractable, I treat scn_{nt} as exogenous as well.

and price difference p_{nt} of

$$\rho_{nt}^{A,cn}(m,p_{nt}) = (1-\bar{\rho}_{nt}^{H})\frac{\exp\{\delta_{nt}^{A}(p_{nt})+m\}}{1+\exp\{\delta_{nt}^{A}(p_{nt})+m\}}.$$

When hotels' capacity constraints do not bind, the marginal customer has the choice between all three options, and the expected Airbnb market share for posterior mean m and price difference p_{nt} is given by the nested logit choice probability

$$\rho_{nt}^{A,\neg cn}(m,p_{nt}) = \frac{\exp\left\{\frac{\delta_{nt}^{A}(p_{nt})+m}{1-\zeta}\right\}\left(\exp\left\{\frac{\delta_{nt}^{H}}{1-\zeta}\right\}+\exp\left\{\frac{\delta_{nt}^{A}(p_{nt})+m}{1-\zeta}\right\}\right)^{-\zeta}}{1+\left(\exp\left\{\frac{\delta_{nt}^{H}}{1-\zeta}\right\}+\exp\left\{\frac{\delta_{nt}^{A}(p_{nt})+m}{1-\zeta}\right\}\right)^{1-\zeta}}$$

resulting from the full demand system specified above.

Market Shares and Platform Profits. The overall expected Airbnb market share as a function of posterior means and the price difference is then given by $\rho_{nt}^{A}(m,p_{nt}) = scn_{nt}\rho_{nt}^{A,cn}(m,p_{nt}) + (1-scn_{nt})\rho_{nt}^{A,\neg cn}(m,p_{nt})$. Recall that for fixed prices and host effort, the Airbnb market share is proportional to Airbnb profits. Hence, for fixed p_{nt}, $\rho_{nt}^{A}(m)$ is the function that is depicted in the top panels of Figure 2, the objective function of the Bayesian persuasion problem conditional on posterior means.

The aggregate market share for Airbnb is then given by

$$\begin{aligned}\rho^{A} &= \int_{\underline{\omega}_{nt}}^{\infty}\int_{\mathcal{S}}\pi_{nt}(s|\omega)\rho_{nt}^{A}\left(\mathbb{E}_{\mu^{s}}[\omega]\right)ds\,dF(\omega;e_{nt})\\ &= \int_{\underline{\omega}_{nt}}^{\infty}\rho_{nt}^{A}(m)dG_{nt}(m). \end{aligned} \tag{6}$$

The first expression explicitly states the double integral over qualities ω and conditional ratings $s|\omega$ resulting from the information structure π_{nt}. The second expression, instead, exploits the shortcut that the expected market share of an information structure depends on the receiver's (customers') posterior beliefs only through the posterior means $m = \mathbb{E}_{\mu_{s}}[\omega]$ that it induces, and integrates directly with respect to the distribution G_{nt} of posterior means.[36] Consequently, platform profits can also be

[36]Along the same lines, the expression for the aggregate hotel market share is given by

$$\begin{aligned}\rho_{nt}^{H} = (1-scn_{nt})\Bigg(&F(\underline{\omega}_{nt};e_{nt})\frac{\exp\{\delta^{H}\}}{1+\exp\{\delta^{H}\}} \\ &+\int_{\underline{\omega}_{nt}}^{\infty}\frac{\exp\left\{\frac{\delta_{nt}^{H}}{1-\zeta}\right\}\left(\exp\left\{\frac{\delta_{nt}^{H}}{1-\zeta}\right\}+\exp\left\{\frac{\delta_{nt}^{A}+m}{1-\zeta}\right\}\right)^{-\zeta}}{1+\left(\exp\left\{\frac{\delta_{nt}^{H}}{1-\zeta}\right\}+\exp\left\{\frac{\delta_{nt}^{A}+m}{1-\zeta}\right\}\right)^{1-\zeta}}dG_{nt}(m)\Bigg)+scn_{nt}\bar{\rho}_{nt}^{H}.\end{aligned} \tag{7}$$

expressed in terms of posterior means only. That is,

$$\Pi_{nt}^{A} = \underbrace{\tau(p_{nt}^{H} + p_{nt})}_{\substack{\text{Airbnb's share of}\\ \text{transaction price}}} \underbrace{\int_{\underline{\omega}_{nt}}^{\infty} \rho_{nt}^{A}(m, p_{nt}) dG_{nt}(m)}_{\substack{\text{expected Airbnb}\\ \text{market share}}}. \tag{8}$$

This particular case of a Bayesian persuasion problem with such an objective function has been studied extensively in the theoretical literature. Any information structure π_{nt} is a garbling of the fully revealing information structure that generates a distinct signal realization in each state ω and thereby induces a distribution of posterior means equal to the prior distribution F. Hence, given the prior F, a distribution of posterior means G_{nt} is induced by some information structure if and only if F is a mean-preserving spread of G_{nt} (Blackwell, 1953; Gentzkow and Kamenica, 2016). As it is common practice in this literature, I can write down the platform's optimization problem directly over the distribution of posterior means, with F being a mean-preserving spread of G_{nt} as an additional constraint.

Host Objective. Assuming additive separability of the effort and participation costs, the representative host's objective function is given by

$$\begin{aligned}\Pi_{nt}^{RH} &= \underbrace{(1-\tau)(p_{nt}^{H} + p_{nt})}_{\substack{\text{hosts' share of}\\ \text{transaction price}}} \underbrace{\int_{\underline{\omega}_{nt}}^{\infty} \int_{S} \pi_{nt}(s|\omega) \rho_{nt}^{A} \left(\mathbb{E}_{\mu^{s}}[\omega], p_{nt}\right) ds dF(\omega; e_{nt})}_{\substack{\text{expected Airbnb}\\ \text{market share}}} \\ &\quad - \underbrace{\int_{\underline{\omega}_{nt}}^{\infty} T_{nt} dF(\omega; e_{nt})}_{\substack{\text{listing}\\ \text{costs}}} - \underbrace{k_{nt} e_{nt}^{2}}_{\substack{\text{effort}\\ \text{costs}}} \\ &= (1-\tau)(p_{nt}^{H} + p_{nt}) \int_{\underline{\omega}_{nt}}^{\infty} \rho_{nt}^{A} \left(m, p_{nt}\right) dG_{nt}(m) - \int_{\underline{\omega}_{nt}}^{\infty} T_{nt} dF(\omega; e_{nt}) - k_{nt} e_{nt}^{2}.\end{aligned} \tag{9}$$

Recall that the representative host has two choice variables, the cutoff quality for listing $\underline{\omega}$ and the effort level e, which both will be determined by the receptive first-order conditions.

For given prices, the host's first order condition for the optimal choice of the cutoff quality $\underline{\omega}$ is independent of his effort choice and depends on signals and customers' belief structure only at $\underline{\omega}$. If all players correctly anticipate full revelation around this threshold, the host's optimal cutoff quality is given by

$$\underline{\omega}_{nt} = (\rho^A_{nt})^{-1}\left(\frac{T_{nt}}{(1-\tau)(p^H_{nt}+p_{nt})}\right). \tag{10}$$

The host's first order condition for the profit-maximizing effort choice reads, after some rearrangements following Boleslavsky and Kim (2021) (see Appendix B.1 for the derivations), as

$$\begin{aligned}(1-\tau)(p^H_{nt}+p_{nt})\int_{\mathcal{S}}\rho^A_{nt}\left(\mathbb{E}_{\mu^s}[\omega],p_{nt}\right)\mathbb{E}_{\mu^s}\left[\frac{\partial f(\omega;e_{nt})/\partial e_{nt}}{f_{\underline{\omega}_{nt}}(\omega;e_{nt})}\right]d\tilde{G}_{nt}(\mu^s)\\ = T_{nt}f(\underline{\omega}_{nt};e_{nt})+2k_{nt}e_{nt},\end{aligned} \tag{11}$$

where $\tilde{G}_{nt}$ is the distribution of posterior beliefs μ^s.

Note that in the general case, this first-order condition does depend on posterior beliefs not only through posterior means but also the posterior expectation of the ratio $\frac{\partial f(\omega;e_{nt})/\partial e_{nt}}{f_{\underline{\omega}_{nt}}(\omega;e_{nt})}$ appears. Abstracting from the truncation, this ratio can be interpreted as the score (i.e., the derivative of the log-likelihood function) of the maximum likelihood estimation of the mean of the prior distribution. On a very intuitive level, it tells how strongly one is inclined to infer e from any observed ω.[37] Therefore, on the left-hand side of the first order condition 11, we have the marginal benefits of exerting more effort in the form of an increased booking probability for all listings weighted by the expected value for the customers' statistical estimate of hosts' effort, and on the right-hand side the marginal costs.

Thanks to the normality assumption of the prior, the ratio $\frac{\partial f(\omega;e_{nt})/\partial e_{nt}}{f_{\underline{\omega}_{nt}}(\omega;e_{nt})}$ reduces to a linear function of ω and the first order condition can be restated depending on posterior means only:

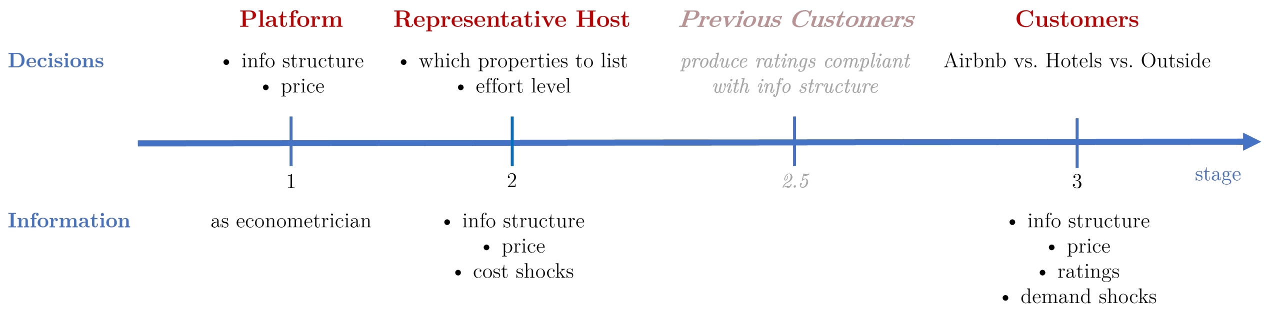
$$\int_{\underline{\omega}}^{\infty}\rho^A_{nt}(m,p_{nt})(m-\nu-e_{nt})dG_{nt}(m) = \frac{\sigma^2[T_{nt}f(\underline{\omega}_{nt},e_{nt})+2k_{nt}e_{nt}]}{(1-\tau)(p^H_{nt}+p_{nt})[1-F(\underline{\omega}_{nt},e_{nt})]}. \tag{12}$$

Thus, thanks to the normality assumption, the platform's optimization problem, including the additional constraint for host effort, depends on posterior beliefs only through means, which keeps the model tractable by allowing me to make use of the results from above mentioned theoretical literature for this special case.[38]

[37] See Boleslavsky and Kim (2021) for further discussion on the role of this ratio.

[38] This is the only place where the normality assumption is needed for tractability of the model. A similar model with exogenous host effort can be solved (and with some parametric assumptions potentially estimated) for an arbitrary prior distribution.

Figure 7: Model Overview and Timing



Timing. The timing of the game is summarized in Figure 7. The game unfolds in three stages. In the first stage, the platform designs and publicly reveals the information structure π_{nt} and chooses a uniform price difference p_{nt}. To keep the complex Bayesian persuasion model tractable, I assume that the platform makes these choices supposing that all unobservable cost and demand shocks are zero (i.e., $\varepsilon^T_{nt} = \xi^A_{nt} = \xi^H_{nt} = 0$).[39] In the second stage, the representative host observes price, information structure, and cost shock and decides how much costly effort to exert and which properties to list. Also here, for tractability I assume that he makes these choices supposing that demand shocks equal zero (i.e., $\xi^A_{nt} = \xi^H_{nt} = 0$). Between the second and the third stage, 'previous customers' produce ratings compliant with π_{nt}. In the third stage, customers observe information structure, price, ratings, and demand shocks and decide between Airbnb, hotels, and the outside option.

All agents are risk neutral and maximize their expected profits or expected utility. Observable quality characteristics of the platform and hotels and the hotel price are assumed to be exogenous and known by all agents.

Equilibrium Definition. I solve for the Perfect Bayesian Equilibrium of this game in each market nt. An equilibrium consists of a price (difference) p_{nt}, an information structure π_{nt}, representative host's effort e_{nt} and cutoff quality $\underline{\omega}_{nt}$ (for each possible signal and price), a customers' posterior belief system $M_{nt} = \{\mu_s\}_{s \in \mathcal{S}}$, and market shares ρ^A_{nt}, ρ^H_{nt}, which satisfy the following properties: (i) given any price p_{nt}, information structure π_{nt}, and customers' belief system M_{nt}, the host's effort e_{nt} and cutoff quality $\underline{\omega}_{nt}$ maximize his expected profit Π^{RH}_{nt}; (ii) customers' belief system M_{nt} is Bayes-consistent with the host's effort e_{nt}, cutoff quality $\underline{\omega}_{nt}$, and the information structure π_{nt}, and market shares ρ^A_{nt}, ρ^H_{nt} result from optimal choices

[39]The solution of a Bayesian persuasion problem taking expectations over all these dimensions would be beyond the capabilities of the existing theoretical literature.

given this belief system; and (iii) price p_{nt} and information structure π_{nt} maximize the platform's expected profit Π^A_{nt}.

Platform Problem with Equilibrium Conditions as Constraints. Finding the equilibrium of this game can be reduced to the solution of the platform's profit maximization problem of choosing a price difference p_{nt}, a targeted host effort level e_{nt}, and an implied distribution of posterior means $G_{nt}(m)$ with the equilibrium conditions as constraints. As derived above, these constraints are (i) that the prior distribution needs to be a mean-preserving spread of the distribution of posterior means (which essentially means that customers update their beliefs according to Bayes' rule) and (ii) that the targeted effort level e_{nt} satisfies the representative host's first order condition. In summary, we have

$$\max_{p_{nt}, e_{nt}, G_{nt}} \tau(p^H_{nt} + p_{nt}) \int_{\underline{\omega}_{nt}}^{\infty} \rho^A_{nt}(m, p_{nt}) dG_{nt}(m) \tag{13}$$

$$\text{s.t. } F_{\underline{\omega}_{nt}}(\cdot; e_{nt}) \text{ is a mean preserving spread of } G_{nt}$$

$$\int_{\underline{\omega}_{nt}}^{\infty} \rho^A_{nt}(m, p_{nt})(m - \nu - e_{nt}) dG_{nt}(m) = \frac{\sigma^2[T_{nt} f(\underline{\omega}_{nt}; e_{nt}) + 2k_{nt}e_{nt}]}{(1-\tau)(p^H_{nt} + p_{nt})[1 - F(\underline{\omega}_{nt}; e_{nt})]}.$$

Solution Algorithm. I exploit recent theoretical results from the duality approach to Bayesian persuasion (Dworczak and Martini, 2019; Dworczak and Kolotilin, 2019)[40] to derive a novel numerical solution algorithm for the platform's maximization problem. A similar algorithm may be applicable more broadly for the numerical solution of Bayesian persuasion problems with additional constraints. Detailed derivations of the algorithm are provided in Appendix B.2.

In summary, I apply backward induction optimization with three nested loops. In the outer loop, I maximize the objective function over price difference and target effort. This means that in the middle loop, for a fixed price difference and target effort, only the profit-maximizing information structure remains to be found, resulting in a Bayesian persuasion problem with moral hazard à la Boleslavsky and Kim (2021). Results from said duality theory allow me to solve this constrained Bayesian persuasion problem by repeatedly solving unconstrained Bayesian persuasion

[40] The price-theoretic duality approach to Bayesian persuasion shows an analogy between the sender's information design problem and finding Walrasian equilibria of a 'persuasion economy'. In the dual problem, the sender acts as a consumer purchasing posterior beliefs at certain prices using the prior distribution as an endowment. The Bayesian updating constraint translates into a single firm having the technology to garble the state. This approach yields a tractable solution method for persuasion problems in which the sender's utility depends only on the expected state by finding the equilibrium price function of the dual problem (Dworczak and Martini, 2019). Moreover, this approach generalizes to Bayesian persuasion problems with additional constraints (Dworczak and Kolotilin, 2019).

problems and minimizing an objective over a shadow-price-like parameter. This minimization is done in the middle loop, and the unconstrained problems are solved in the inner loop.

Solution for S-shaped Objective Function. In the inner loop of the algorithm, I repeatedly solve an unconstrained Bayesian persuasion problem with a continuous state space (i.e., $\mathbb{R}$, the unobserved Airbnb quality), a continuous action space of the receiver (i.e., $[0,1]$, the share of customers choosing the platform), and an S-shaped posterior-means objective function. Recall from the derivation in Section 2 that this objective function corresponds to the cumulative distribution function of customers' reservation values for choosing Airbnb. The S-shape means the distribution is unimodal, with more reservation values in an intermediate range than extremely low or high ones. Examples of demand systems that yield such an S-shaped objective function are logit and nested logit models.

The objective function in the inner loop of my solution algorithm is derived from a more involved demand system accounting for hotels' capacity constraints and is slightly perturbed due to the effort constraint, but it turns out that for all empirically relevant parameter values, the S-shape of this function is preserved.

Therefore, any interior solution[41] of this unconstrained problem with an S-shaped objective function in the inner loop still follows the *upper-censoring* structure, where the cutoff quality ω^*_{nt} is defined according to equation 1. To tie this model prediction to the data, I interpret the pooling signal as 5-star ratings and the more granular revelation of lower qualities as 1–4.5 stars.

5. Estimation

Overview. I estimate the empirical model specified in the previous section by the Generalized Method of Moments (GMM). The empirical specification includes 16 parameters to be estimated: κ_0 and κ_1 defining hosts' effort costs in equation 2, η_0 and η_1 defining hosts' participation cost in equation 3, and the demand parameters specified in equations 4 and 5. The latter are the price coefficient α, coefficients for tourist POIs (ψ), and observed Airbnb and hotel quality characteristics ($\boldsymbol{\beta}^A$ and $\boldsymbol{\beta}^H$), as well as the nesting parameter ζ and mean ν and standard deviation σ of

[41] If

$$\rho^A_{nt}\left(\mathbb{E}_{F_{\underline{\omega}_{nt}}}[\omega]\right) > \rho^A_{nt}(\underline{\omega}) + \left(\mathbb{E}_{F_{\underline{\omega}_{nt}}}[\omega] - \underline{\omega}\right)(\rho^A_{nt})'\left(\mathbb{E}_{F_{\underline{\omega}_{nt}}}[\omega]\right),$$

we have full pooling, i.e., revealing noting about ω, as a boundary solution.

the unobserved Airbnb quality distribution. I treat observable Airbnb and hotel characteristics, the number of tourist points of interest, wages and long-term rents, the share of hotel compression nights, and hotel capacity as exogenous. The estimation targets 24 moments specified below, a combination of orthogonality conditions and levels.

A main identifying assumption is that observed market outcomes (demand, Airbnb prices, share of 5-star ratings) result from the platform choosing an information structure and price optimally in each market, responding to relevant market characteristics (recall the comparative statics in Figure 3), and customers rationally updating their beliefs accordingly. I exploit cross-market variation of $N = 112$ observations from 56 geographical regions over two years, as specified in Section 3.2. In each step of the GMM minimization routine, I solve my model for the platform's optimal choices in all markets.[42]

Recall that the main interest for the counterfactual experiments is to recover the distribution of unobserved Airbnb quality. For this purpose, I assume that the effortless distribution of unobserved quality of potential Airbnb listings, parametrized by ν and σ, is the same across all markets.[43]

Demand Inversion. Since the aggregate market shares (equations 6 and 7) are given as integrals over unobserved Airbnb quality, an analytical inversion of the resulting demand system to recover the structural error terms ξ is not feasible. The problem is reminiscent of the inversion of random coefficient demand systems à la Berry, Levinsohn, and Pakes (1995) (BLP) with the difference that here we integrate over heterogeneous qualities of the supplied service instead of over heterogeneous consumer tastes. Since the sufficient conditions for a unique solution of the demand inversion provided by Berry (1994) (i.e., ρ^A_{nt} is differentiable with respect to δ^A_{nt} and δ^H_{nt} with $\frac{\partial \rho^A_{nt}}{\partial \delta^A_{nt}} > 0$ and $\frac{\partial \rho^A_{nt}}{\partial \delta^H_{nt}} < 0$, and ρ^H_{nt} is differentiable with respect to δ^H_{nt} and δ^A_{nt} with $\frac{\partial \rho^H_{nt}}{\partial \delta^H_{nt}} > 0$ and $\frac{\partial \rho^H_{nt}}{\partial \delta^A_{nt}} < 0$) are satisfied, the demand system can be inverted numerically to recover the structural error terms ξ just as in BLP.

[42]The estimation is carried out in the programming language Julia (Bezanson, Edelman, Karpinski, and Shah, 2017) on a 32-core ScienceCloud server at the University of Zurich, using the Simulated Annealing algorithm (Goffe, Ferrier, and Rogers, 1994; Goffe, 1996) to find the global minimum of the GMM objective function. This probabilistic global optimization algorithm takes over 200,000 objective function evaluations to converge. Only the high computational performance of the Julia language (one objective function evaluation solving the model and inverting market shares for 112 markets in less than one second) allows me to estimate this complex model with endogenous host effort by reducing computation time by an order of magnitude compared to a MATLAB implementation of the same algorithm.

[43]Similar assumptions have been made by Bertomeu, Ma, and Marinovic (2020) and Bertomeu, Marinovic, Terry, and Varas (2022) to estimate models of information disclosure by managers.

Note that even though this inversion transforms the demand system into a linear system, I do not estimate the remaining demand parameters entering the observable utility components δ in linear regressions outside the main GMM estimation routine, as it is commonly done in BLP models. This is because, in my model, the observable utility components δ also enter the non-linear part (the distribution G with respect to which the market shares are integrated) as the optimal information structure depends on them. Therefore, all 16 parameters of the model are estimated simultaneously in the main GMM routine. The high computational cost of each additional parameter in such a non-linear estimation explains my parsimonious parametrization of the demand model with a relatively small number of additional observable quality characteristics as control variables.

Inversion of Host Participation. To disentangle the impact of different factors on the quality distribution and observed outcomes, it is convenient also to recover ε^T, the structural error term in hosts' opportunity cost for participation as introduced in equation 3. For this purpose, in each step of the GMM estimation routine and for each market, I construct an empirical equivalent of $\underline{\omega}$ by assuming that Airbnb's maximum attainable market share after hosts' participation choice, $1 - F(\underline{\omega}; 0)$, is equal to the empirically observed Airbnb capacity divided by the market size.[44] Given this empirical value of $\underline{\omega}$, I invert equation 10 to obtain an empirical equivalent of the opportunity cost T, which then allows me to calculate the structural error term ε^T in equation 3.

Moments. Table 2 summarizes the 24 targeted moments, a combination of orthogonality conditions and levels.[45]

The first rows contain the usual moments for nested logit demand estimation. The error terms ξ are set orthogonal to all the exogenous covariates of equations 4 and 5 and excluded instruments for prices and inside market shares. In particular, I instrument for Airbnb prices with the daily long-term rent as a supply shifter,

[44]This assumption essentially means that for $\delta^A = \infty$, the demand model would yield a predicted Airbnb market share equal to the current Airbnb capacity (keeping all the remaining moving parts on the supply side constant). Hence, it is abstracting from Airbnb's supply being elastic with respect to prices or anticipated demand. While this would be a quite strong assumption at a daily level, with my yearly aggregation, it is reasonable. The elastic portion of Airbnb supply stems mainly from occasional peer-to-peer hosts (instead of professional investors who offer their apartments all year round). Those peer-to-peer hosts might be willing to host someone in their place for a couple of days a year when the price is high enough, but probably not 365 days, even if the price was always this high.

[45]For the construction of the GMM objective function, orthogonality moments derived from the same equation are weighted by the homoskedasticity-optimal weighting matrix. In addition, across the equations, moments are rescaled to a comparable order of magnitude.

Table 2: Target Moments for GMM Estimation

orthog.	$\hat{\xi}^A \perp \{\ \mathbb{1}_u,\ \mathbb{1}_{nu},\ \boldsymbol{q}^A,\ \log(poi),\ \text{rent},\ q^H\ \}$
conditions	$\hat{\xi}^H \perp \{\ \mathbb{1}_u,\ \mathbb{1}_{nu},\ \boldsymbol{q}^H,\ \log(poi),\ \text{unavailable land},\ q^A\ \}$
	$\hat{p} - p \perp \{\ \mathbb{1},\ p^H,\ \text{rent},\ \text{wage}\ \}$
	$\widehat{\varepsilon^T} \perp \{\ \mathbb{1},\ \text{rent},\ \text{wage}\ \}$
levels	mean squared error of predicted share of 5-star ratings = 0
	occupancy non-5-star listings/occupancy 5-star listings = prediction

Notes: The indicator functions $\mathbb{1}_u$ and $\mathbb{1}_{nu}$ identify the groups of urban and non-urban travel destinations, as classified in Table A.1. The remaining variables are defined in equations 2, 3, 4, and 5. Hats indicate model predictions. Excluded instruments are highlighted in red.

motivated by my model of host participation (equation 3), and for hotel prices with the share of unavailable land (Lutz and Sand, 2019) in the respective market. In the latter, I follow Farronato and Fradkin (2018) who find that lack of suitable building ground is a significant driver restricting the supply of additional hotel rooms, therefore leading to higher hotel prices. Moreover, as standard in the literature on nested logit demand estimation for the identification of the nesting parameter ζ, I use BLP-style instruments for the inside market shares. More precisely, I use the observable quality characteristic of the other option with the strongest first stage.[46]

As shown in the third row, I chose to set the error term of the predicted price difference orthogonal to a constant and the hotel price because the price differences predicted by the model are generally well approximated by a linear function of hotel prices (see Figure 8), as well as to the covariates of the cost equations that shift the prior quality distribution. Along the same lines, the error term for host opportunity costs of participation ε^T is set orthogonal to all the covariates of equations 2, 3 for host effort and participation costs, which can affect the prior quality distribution directly.

In addition to these orthogonality moments, I target two moments in levels. First, I set the mean squared prediction error of the share of 5-star ratings, calculated as

$$\widehat{share5star}_{nt} = \frac{1 - F(\omega^*_{nt}; e_{nt})}{1 - F(\underline{\omega}_{nt}; e_{nt})},$$

equal zero. Since in the light of the research question, the share of 5-star ratings is a key outcome of my model, which should be matched well in every market, in this way, I put increased emphasis on it instead of just matching the mean across

[46]First stage results for all the excluded instruments are displayed in Table A.2.

markets. Moreover, I target the expected occupancy ratio of non-5-star and 5-star listings, whose model prediction is calculated as

$$\widehat{occ\_rat}_{nt} = \frac{\frac{\int_{\underline{\omega}_{nt}}^{\omega^*_{nt}} \rho^A_{nt}(m) dG_{nt}(m)}{F(\omega^*_{nt}; e_{nt}) - F(\underline{\omega}_{nt}; e_{nt})}}{\rho^A_{nt}(\tilde{\omega}_{nt})}.$$

Identification. The complexity of the model precludes a formal proof of identification. However, the targeted moments are picked such that there is a clear intuition of how, ceteris paribus, they are informative about specific parameters.

Identification of the observable market and quality characteristic and price parameters in equations 4 and 5 follows the usual logic in the demand estimation literature via the orthogonality conditions between structural error terms ξ and covariates as well as instruments (first two rows of Table 2). Note that the set of parameters identified in this way also includes the mean of the effortless unobserved quality distribution ν, which can also be interpreted as a fixed effect for Airbnb. I use the instruments discussed above to alleviate the usual endogeneity concerns about equilibrium outcomes.

In combination with the assumption on the maximum attainable Airbnb market share discussed above, which allows me to construct empirical equivalents of $\underline{\omega}_{nt}$, the orthogonality conditions between the error term ε^T and the covariates in the host participation cost equation 3 (row 4 of Table 2) pin down the respective parameters η_0 and η_1.

The occupancy ratio of non-5-star and 5-star listings $\widehat{occ\_rat}_{nt}$ is informative about the standard deviation of the unobserved quality distribution σ. Intuitively, if σ is larger and the distribution is more dispersed, the difference in the expected choice probabilities for non-5-star and 5-star listings increases, and $\widehat{occ\_rat}_{nt}$ decreases. As an essential identifying assumption is that σ is constant across markets, I target a simple average of this occupancy ratio across markets. Empirically, the observed variance of this ratio across markets is relatively small and does not appear to correlate systematically with any other influential market characteristics such as prices (see Figure A.1). While these empirical patterns cannot prove that the identifying assumption is accurate, they do not provide evidence against it.

Finally, we are left with the parameters κ_0 and κ_1, which pin down the marginal effort cost as defined in equation 2. This cost eventually defines the effort level e_{nt}, which shifts the mean of the prior distribution of Airbnb quality and determines its ex-ante attractiveness vis-a-vis the alternatives. The numerical solutions show that the model features an inherent substitutability between higher prices and less information (i.e., pooling of more qualities in the 5-star rating) for the platform's response to an increased attractiveness ex-ante; recall the intuitive discussion in Section 2. Given this substitutability, for a fixed shape of the objective function $\rho^A_{nt}(m)$, the combination of Airbnb price level and the share of 5-star ratings allows me to infer the mean of the prior distribution and the corresponding marginal effort cost.

6. Results

Parameter Estimates and Interpretation. Table 3 displays the estimated parameter values and standard errors. To facilitate the interpretation, the parameters of the demand equations can be transformed into monetary values by dividing them by the price coefficient α.

The main parameter of interest for the counterfactuals, the standard deviation of the unobserved Airbnb quality distribution of potential listings σ (before the left-truncation through hosts' participation choice), has a monetary equivalent of US$94.64, corresponding to 68% of the average price paid for one night booked on Airbnb. Hence, my estimates suggest a substantial degree of unobserved quality heterogeneity of potential listings, which could completely overturn the surplus derived from ex-ante observable characteristics. The standard deviation of unobserved quality of *actual* listings (after the left-truncation through hosts' participation choice) ranges between US$27.16 and US$45.50 across different markets, with an average of US$34.66. Hence, the information provided through the rating system with informative ratings at the bottom of the distribution allows the platform to filter out the worst qualities reducing the dispersion in the unobserved quality distribution by up to 70%.

The positive (and statistically significant) estimates of $\boldsymbol{\beta}^H$ confirm that the attractiveness of hotels increases in their quality, measured by the STR classification and booking.com ratings. Even though quite noisy, the estimates of $\boldsymbol{\beta}^A$ suggest that one-bedroom entire homes (the omitted category) are the most attractive Airbnb listing type. Urban destinations appear to be slightly more attractive than non-urban ones ($\gamma_u > \gamma_{nu}$), and the estimate of ψ shows that more tourist points of interest in a market increase its attractiveness.

On the supply side, hosts' average opportunity cost of listing their properties on Airbnb (η_0) is US$35.14 per night and appears to be mostly independent of the

Table 3: Parameter Estimates

		Parameter Estimate	Standard Error
Demand Coefficients	Price (α) [4,5]	0.01353	0.00137
	FE urban destination (γ_u) [4,5]	0.54	0.22
	FE non-urban destination (γ_{nu}) [4,5]	0.51	0.12
	Log OSM tourist POIs (ψ) [4,5]	0.32	0.08
	Share 'private room' (β_1^A) [4]	-7.82	1.70
	Share 'entire home' studio (β_2^A) [4]	-0.73	2.77
	Share 'entire home' 2 bedroom (β_3^A) [4]	-5.69	2.20
	Avg. bookig.com rating (β_1^H) [5]	1.52	0.76
	Share $\geq$ Upper Midscale (β_2^H) [5]	0.81	0.31
	Nesting parameter (ζ) [4,5]	0.232	0.573
Unobs. Airbnb Quality	Mean (ν)	-1.51	0.39
	Std. dev. (σ)	1.28	0.27
Effort Costs	Constant (κ_0) [2]	55.12	19.41
	Wage coefficient (κ_1) [2]	0.12	2.02
Listing Costs	Constant (η_0) [3]	35.14	1.47
	Rent coefficient (η_1) [3]	-0.007	0.09

Notes: GMM estimates and standard errors of the model parameters. The numbers in brackets refer to the equations the parameters are defined in. Standard errors are calculated based on the usual analytical formula for GMM standard errors, based on numerical derivatives of the objective function at the optimum.

long-term rent level. Though very noisy, the estimate of κ_1 suggests that a one-dollar increase in average wages increases hosts' opportunity cost for effort provision by 12 cents.

Demand Elasticities. Table 4 shows the resulting demand elasticities averaged across markets. Despite the relatively coarse demand estimation strategy relying mainly on the cross-sectional variation between rather large geographical regions (recall the discussion at the end of Section 3.2), the estimated elasticities range in a comparable order of magnitude to those of Farronato and Fradkin (2022) and Schaefer and Tran (2020), who focus on demand estimation exclusively with much more granular data.

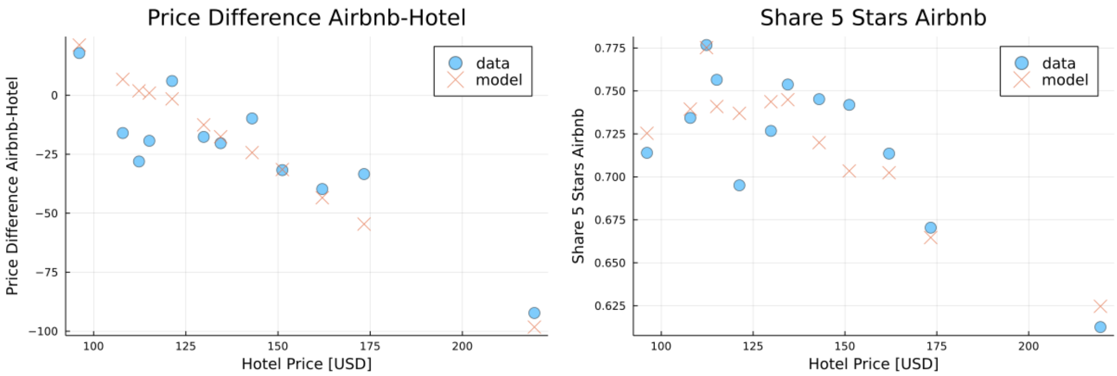
Model Fit. Figure 8 repeats the binned scatter plots of Airbnb 5-star ratings and the Airbnb-hotel price difference with the average hotel prices on the horizontal axes from Figure 5 and overlays it with the model predictions for the estimated parameters. The predicted values fit very closely the empirical cross-market patterns

Table 4: Average Demand Elasticities

	Airbnb	Hotel	Outside
Airbnb	-1.94	0.90	-
Hotel	0.10	-1.55	-
Outside	0.04	0.39	-

Notes: Nested logit demand elasticities for non-compression nights (i.e., nights in which customers have the choice between all three options) averaged across 112 markets.

Figure 8: Model Fit



Notes: The graphs show binned scatter plots of empirical realizations (blue circles) and fitted model predictions (red crosses) of the Airbnb-hotel price difference (left panel) and the share of Airbnb 5-star ratings (right panel) with the average hotel price across markets on the horizontal axes.

of these two crucial targeted moments; that is, they match well both the linear negative relationship between hotel prices and price difference and the hump-shaped relationship between hotel prices and the share of Airbnb 5-star ratings, which were discussed in Section 2.

Estimation Results without Demand Instruments. Appendix Table A.3 shows parameter estimates and standard errors in an alternative specification without demand instruments. As expected, given the usual endogeneity concerns regarding prices and inside market shares, without instruments, the price parameter α and nesting parameter ζ are underestimated compared to the specification with instruments. As a consequence, also the resulting demand elasticities, which are displayed in Table A.4, are slightly attenuated. However, the differences between the two specifications are relatively small, which shows the substantial role the supply-side restrictions (i.e., the moments concerning prices and the share of 5-star ratings) play in pinning down α and ζ. All the remaining parameters are not affected substantially.

7. Counterfactual Experiments

Overview. With the estimated model parameters at hand, I can simulate counterfactual scenarios with different rating system designs. In particular, I will consider scenarios with no ratings at all and perfectly revealing ratings to quantify welfare consequences resulting from the platform's strategic rating system design, as well as some intermediate cases. Moreover, my model allows me to disentangle to which extent welfare changes can be attributed to the role of ratings to incentivize host effort and to which extent to their role of informing customers. Finally, I will evaluate the heterogeneous impact of strategic rating design on Airbnb, host, and hotel profits, depending on Airbnb listing quality. Appendix C contains additional counterfactual exercises with platform-optimal and consumer-optimal binary ratings.

Market outcomes and welfare measures under different counterfactual scenarios averaged across markets are summarized in Table 5. The baseline outcomes in column 1 are those resulting from platform-optimal upper-censoring ratings, which by assumption, coincide with the empirically observed situation.

Outcome Measures. All quantities are converted into monetary terms by division by the estimated price coefficient α. In the case of platform-optimal ratings, Airbnb and hotel market shares are calculated as stated in equations 6 and 7. In the counterfactuals with fully revealing ratings, integration is done with respect to the prior distribution of unobserved Airbnb quality, and in the cases without ratings, the functions are evaluated at the mean of this prior distribution. The formulas for these calculations are spelled out in Appendix B.3. Airbnb and host profits per potential listing are calculated according to equations 8 and 9 in the baseline, with the same adjustments to the integrals in the counterfactual scenarios as for the market shares. Profits per actual listing are obtained by dividing profits per potential listing by the mass of listings the host actually decides to offer on the platform, that is $1 - F(\underline{\omega}; e)$.

Since consumer surplus in discrete choice models is only defined up to an additive constant, as a measure of the value of Airbnb for customers, I use the Compensating Variation (CV) (Small and Rosen, 1981) relative to a counterfactual where I remove Airbnb completely from the choice set.[47] Hence, the second-to-last row in Table 5 shows the dollar amount that would need to be paid to every potential customer in a

[47] Aggregated over the 56 geographical markets and the years 2018 and 2019, I estimate a CV for Airbnb of US\$12.3B. As a point of reference, for New York City, I estimate a CV for Airbnb of US\$288.9M in 2018 (4.5M booked nights) and US\$247.6M in 2019 (7.1M booked nights), which appears reasonable in comparison to the estimate of US\$141.3M in 2014 (1.8M booked nights) for the same market by Farronato and Fradkin (2022).

market ex-ante to make up for the expected utility loss from removing Airbnb from the choice set, calculated as

$$CV = \frac{1}{\alpha}\left(\mathbb{E}[\max\{U^A, U^H, U^0\}] - \mathbb{E}[\max\{U^H, U^0\}]\right), \tag{14}$$

where expectations are taken both over the idiosyncratic taste shocks and posterior believed Airbnb quality. The exact formulas for how these expectations are calculated for the demand system accounting for hotels' capacity constraints and different choice sets for the marginal Airbnb customer in compression and non-compression nights are provided in Appendix B.3.

Note that with rational Bayesian updating of customer beliefs, the ex-ante expected CV is equal to the average ex-post realized surplus for any information structure. A correction term for the difference between anticipated and experienced quality, as suggested by Train (2015), would cancel out when integrating over the quality distribution.

The value of Airbnb per booked night in the last row of Table 5 is obtained by dividing the CV from equation 14 by the Airbnb market share. Airbnb's average baseline value per booked night is estimated as 79% of the average Airbnb price.

Informative Roles of Ratings. The informative role of customer ratings can be subdivided into two parts. First, they allow filtering out the worst qualities from joining the platform. Second, they enable consumers to learn more about the exact quality of those properties that are actually listed on the platform. The counterfactual exercise without any rating system reveals how vital the first role is for the functioning of Airbnb and that the platform-optimal upper-censoring ratings do a good job in this regard, as they fully reveal the quality for the worst qualities and those around the participation cutoff $\underline{\omega}$.

No Ratings. In a counterfactual where I completely remove the customer rating system (column 2 of Table 5), both these roles get shut down. In this case, the platform is not able to operate at all. Without informative ratings at the bottom to filter out the worst qualities, potentially everyone could join the platform ($\underline{\omega} = -\infty$), and no additional host effort can be incentivized ($e = 0$). Therefore, consumers' posterior belief about unobserved quality corresponds to the mean ν of the untruncated prior distribution $F(\cdot; 0)$. This believed quality is so low that, independent of the price the platform sets, the probability that customers choose the platform over the other options is so low that expected host profits are well below the fixed costs T for listing a property, and no one would want to offer their properties on the platform. Note

Table 5: Counterfactuals

	(1) **Baseline** (Current platform-optimal upper-censoring ratings)	(2) **No Ratings**	(3) **No Ratings** (Price & participation fixed as in baseline)	(4) **Fully Revealing Ratings**	(5) **Disentangling Information and Incentives** (Baseline but effort as with fully revealing ratings)
Airbnb Price [US$]	117.81	-	117.81	118.34	117.81
			0%	+0.45%	0%
Effort [US$]	0.95	0	0	1.02	1.02
		-100%	-100%	+7.64%	+7.64%
Airbnb Market Share	0.0294	0	0.0284	0.0289	0.0294
		-100%	-3.27%	-1.46%	+0.28%
Hotel Market Share	0.2147	0.2231	0.2150	0.2149	0.2147
		+3.88%	+0.11%	+0.07%	-0.01%
Platform Profit per	0.57	0	0.55	0.56	0.57
***Potential* Listing [US$]**		-100%	-3.31%	-0.89%	+0.29%
Platform Profit per	9.28	-	9.27	9.21	9.28
***Actual* Listing [US$]**			-0.12%	-0.71%	+0.01%
Host Profit per	0.91	0	0.89	0.88	0.91
***Potential* Listing [US$]**		-100%	-2.07%	-3.38%	+0.03%
Host Profit per	13.42	-	13.51	13.05	13.40
***Actual* Listing [US$]**			+0.69%	-2.75%	-0.14%
Ex-ante Customer CV	2.76	0	2.61	2.84	2.77
for Airbnb [US$]		-100%	-5.47%	+2.68%	+0.30%
Airbnb Customer CV	92.60	-	91.04	95.85	92.60
per Booked Night [US$]			-1.68%	+3.52%	+0.01%

Notes: Averages across 112 markets.

that this result is not built into the model specification but is an empirical result for the estimated model parameters.[48]

In column 3, I consider an intermediate case, where I fix the Airbnb prices and hosts' participation decision (i.e., $\underline{\omega}$) at their baseline levels to isolate the impact of the second role of ratings in providing further information about *actually listed* properties. Compared to upper-censoring ratings, now customers cannot identify the worst qualities *among those listed on the platform*. Their posterior belief about the unobserved quality of all listings corresponds to the mean of the truncated prior distribution $F_{\underline{\omega}}(\cdot; 0)$ (i.e., the mean quality of those apartments listed). Airbnb and host profits decrease, respectively, by 3.31% and 2.07%, and the ex-ante value of Airbnb to customers is reduced by 5.47% compared to the status quo with platform-optimal upper-censoring ratings.

Fully Revealing Ratings. While customer and platform interests are aligned to prevent the worst qualities from joining the platform, this is not the case for providing full information about the quality of those properties actually listed. In the counterfactual scenario presented in column 4 of Table 5, which addresses one of the leading research questions of this paper, I replace upper-censoring ratings with fully revealing ratings,[49] but let Airbnb re-optimize prices and hosts adjust their effort and participation decisions. With fully revealing ratings, the distribution of posterior beliefs about quality, which is used to calculate market shares and profits, corresponds to the truncated prior distribution $F_{\underline{\omega}}(\cdot; e)$.

In this scenario, Airbnb increases the price slightly, confirming the substitutability between higher prices and less information outlined in the description of Figure 8. Host effort increases by 7.64%.[50] Airbnb and host profits decrease following a market

[48] This counterfactual also reveals that in the absence of Airbnb, its market share of 2.9% would be divided into a 0.8 percentage points higher hotel market share and a 2.1 percentage points higher market share of the outside good. The larger increase in the outside market share is in line with the findings of Farronato and Fradkin (2022) that a large part of Airbnb's business is concentrated on compression nights when no additional hotel rooms are available anyhow. They estimated that in 2014 62% of Airbnb bookings would not have been hotel bookings in a counterfactual scenario without Airbnb.

[49] While continuous fully revealing ratings will never be implementable practically, they still provide an insightful benchmark, and the 9-point scale of Airbnb ratings (steps of half stars) should be able to approximate the consumer welfare gains from fully revealing ratings quite closely. Wilson (1989) proves that in the general case for a partition with n classes, losses relative to full information are of order $1/n^2$, and Hopenhayn and Saeedi (2022a) find that optimal discrete ratings can approximate full information reasonably well in a similar model of rating system design.

[50] Note that fully revealing ratings generally will not maximize induced host effort. In fact, ex-ante, it is not even clear that fully revealing ratings induce more host effort than upper-censoring ratings. Hence, this should be seen as an empirical result. Maximizing induced host effort would be a Bayesian persuasion problem by itself, with a more complex objective function.

Table 6: Aggregated Welfare Impact of Strategic Rating System Design in Monetary Terms

	2018	**2019**	**Total**
Consumer Surplus [M US$]	-127.44	-160.80	-288.23
Hotel Revenue [M US$]	-29.68	-38.03	-67.71
Airbnb Profit [M US$]	+7.85	+10.06	+17.91
Host Profit [M US$]	+44.89	+58.05	+102.95
'Total Welfare' [M US$]	-104.37	-130.72	-235.09

Notes: Differences in welfare measures between the baseline of platform-optimal ratings (column 1 of Table 5) and the counterfactual of fully revealing ratings (column 4 of Table 5) in monetary terms aggregated over the 56 geographical markets (covering about 56% of Airbnb listings in the US).

share decrease of 1.46%. The ex-ante value of Airbnb for customers increases by 2.68%, and the value per actually booked night increases by 3.52%.

Table 6 presents these differences in welfare measures not as averages but aggregated over the 56 geographical areas (covering about 50% of Airbnb listings in the US) and two years in my sample. Despite being much more subtle than the impact of removing the rating system altogether, the aggregate loss in Airbnb consumer surplus due to their strategic rating system design (i.e., platform-optimal upper-censoring ratings vs. fully revealing ratings) adds up to a sizable figure of US$288.2M. Moreover, overall hotel revenue is harmed by US$67.71M, while Airbnb and host profits grow by US$17.91M and US$102.95M, respectively. Adding up all these numbers also suggests a negative net welfare impact of strategic rating system design.[51]

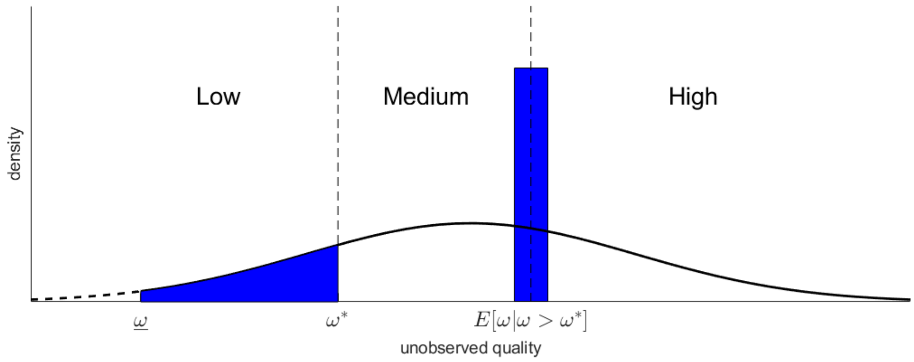
While the increase in Airbnb's profit resulting from the strategic upper-censoring rating design might appear small at first glace (making up only about 0.9% of the profits as predicted by the model), it has to be considered that the Airbnb profit measure in my model does not take into account any costs of operating the platform. Airbnb's financial statements report net losses of US$17M in 2018 and US$674M in 2019 for the worldwide operating company as a whole.[52] The comparison with these actual profit measures reveals that increased profits of US$17.91M from only about half the US market over two years are not negligible at all and may be worth the platform's effort of influencing ratings in said way.

[51]This is within the limitations of the model. For instance, hotel supply is not modeled.
[52]Source: https://www.wsj.com/market-data/quotes/ABNB/financials/annual/income-statement (accessed on October 1, 2022).

Disentangling the Impact of Information and Incentives. The counterfactual with fully revealing ratings has shown that the strategic rating system design harms customers through two channels. First, upper-censoring ratings induce less host effort, shifting the whole quality distribution to the left and decreasing the quality of all properties. Second, some customers make an ex-post wrong decision and choose Airbnb in cases in which the realized quality lies below their reservation value or cannot choose Airbnb when the realized quality lies above their reservation value, but they cannot identify the very best listings. My model allows me to disentangle these two channels. Column 5 of Table 5 presents an intermediate scenario in which I isolate the effort-incentivizing role of ratings. In particular, I keep prices, host participation, and information structure as in the baseline but just adjust the host effort level to the one induced by fully revealing ratings as in column 4. In this scenario, Airbnb's consumer surplus is increased by 0.30%, accounting for only about 1/9 of the overall increase of consumer surplus from fully revealing ratings, leaving about 8/9 to the information channel. This shows that the primary mechanism through which the current Airbnb rating system is disadvantageous for consumers is that it does not let them distinguish 'mediocre' from 'outstanding' listings as they all receive 5 stars. As such, the novel role of ratings in impacting customers' perception of the platform relative to other options, which this paper sheds light on, appears much more impactful than their often focused role of mitigating hosts' moral hazard (conditional on keeping the participation threshold fixed).

Heterogeneous Impact by Airbnb Listing Quality. While Airbnb hosts on average profit from the platform's upper-censoring rating design, not all hosts are actually better off than under fully revealing ratings. There is a redistribution of surplus from high-quality hosts to medium-quality hosts, which exceeds the average effect by almost an order of magnitude. Figure 9 shows the heterogeneous impact on host profits compared to fully revealing ratings (in monetary terms, aggregated over all 56 geographical markets and the years 2018 and 2019), depending on their quality. Low-quality hosts, defined as $\omega < \omega^*_{nt}$, below the threshold for pooling into 5-star ratings, are barely affected (only marginally through changes in price and participation cutoff). Medium quality hosts defined as $\omega \in [\omega^*_{nt}, \tilde{\omega}_{nt}]$ are those that benefit immensely from being pooled into the 5-star rating because the expected quality of the pooling signal is higher than their actual quality and therefore increases the probability that they are booked. Their aggregate profits increase by US$849M, corresponding to 65.9% of aggregated host profits under full information. Also, the Airbnb profit generated from those listings is substantially increased. To the

Figure 9: Heterogeneous Impact by Airbnb Listing Quality



	Total	Low	Medium	High
Airbnb Profit [M US$]	+17.91	-0.54	+161.18	-142.72
Host Profit [M US$]	+102.95	+1.02	+849.24	-747.31
Hotel Revenue [M US$]	-67.71	-2.82	-218.72	+153.83

Notes: Differences in supply-side welfare measures between the baseline of platform-optimal ratings (column 1 of Table 5) and the counterfactual of fully revealing ratings (column 4 of Table 5) in monetary terms aggregated over the 56 geographical markets and years 2018 and 2019, depending on Airbnb listing quality. Low qualities: $\omega < \omega^*_{nt}$; medium qualities: $\omega \in [\omega^*_{nt}, \tilde{\omega}_{nt}]$; high qualities: $\omega > \tilde{\omega}_{nt}$.

contrary, high-quality hosts, defined as $\omega > \tilde{\omega}_{nt}$, are actually harmed by the upper-censoring ratings, as their actual quality is higher than the expected quality induced by the pooling signal. Consequently, their probability of being booked is decreased. Their aggregated profits are US$747.3M lower than under full information. This corresponds to a decrease of 35.2% of the full information profits. Also, the profit Airbnb generates from those listings is lower. However, the losses Airbnb suffers from these listings are outweighed by the additional gains from the medium qualities. This is the case because, with the S-shaped cumulative distribution function of reservation values (or, equivalently, conditional choice probability function) resulting from the estimated demand system, there are more customers with reservation values in the medium range than in the high range, providing some additional intuition why pooling of qualities in this part of the distribution is indeed profit-maximizing.

8. Conclusion

This paper presents a structural estimation of a Bayesian persuasion game embedded in an empirical model of the short-term accommodation market to investigate how

Airbnb's rating system design affects market outcomes and welfare. The theoretical framework highlights the platform's incentives to encourage an excess of rather uninformative 5-star ratings (upper-censoring information structure). My estimates from variation across 56 major travel destinations in the US in the years 2018 and 2019 reveal that a rating system that identifies and allows to exclude the worst qualities is indispensable for the platform's functioning. However, they also suggest that the strategic rating system design reduces the average value of Airbnb to customers by about 3.5% per booked night compared to a counterfactual of fully revealing ratings. Aggregated over the markets and time period studied, this loss amounts to US$288.2M. The platform-optimal upper-censoring ratings negatively affect consumer welfare through two channels, and my model allows me to quantify the relative importance of either one. 1/9 of the consumer welfare loss can be attributed to the pooling of qualities incentivizing less costly host effort. The remaining 8/9 are due to customers' inability to distinguish mediocre from outstanding listings.

While pooling qualities into a 5-star rating harms the highest-quality Airbnb hosts, these losses are more than made up for by gains from listings whose booking probability increases by being included in the pooling. This is because the estimated demand system yields a unimodal distribution of reservation values with more customers in an intermediate range, who can additionally be persuaded, than customers with such high reservation values that they do not consider the platform anymore without being able to identify the very best listings. Overall, fully revealing ratings would reduce aggregated Airbnb profits by US$17.9M and aggregated host profits by US$103M compared to the platform-optimal upper-censoring ratings. This aggregated effect hides, however, that there is a redistribution of surplus from high-quality to medium-quality hosts of almost US$750M.

An essential prerequisite for Airbnb to be able to exploit such a strategic design of the rating system to influence customers' beliefs is that the ratings on the platform are the single only source of information customers have about the unobserved quality of a specific Airbnb listing. If a competitor listed and provided information about the same properties, or if there was an independent platform collecting ratings of Airbnb listings, Airbnb would lose this opportunity. As such, introducing an independent rating platform could be a potential countermeasure if consumer protection authorities got concerned about welfare losses due to strategic rating system design. Hotel booking platforms, where such alternative information sources exist, feature more informative rating systems; recall the distribution of ratings from booking.com in

Figure 1 in the Introduction.

My newly developed conceptual framework for the structural estimation of a numerically solvable Bayesian persuasion problem with an objective function derived from a discrete choice demand system may point the way to study empirically other applications of Bayesian persuasion. It may contribute to bridging the gap between the vast theoretical literature on information design from the last decade and practical applications. The framework could readily be applied to analyze the rating system design of other sharing economy platforms. For instance, the ride-hailing platform Uber appears to follow a similar upper-censoring strategy to Airbnb.[53] Another potential application may be a quantitative welfare analysis of grading inflation in schools and universities, which has been studied in a similar theoretical context by Ostrovsky and Schwarz (2010) and Boleslavsky and Cotton (2015).

[53]See Athey, Castillo, and Chandar (2021) for the distribution of Uber ratings and Appendix D.2 for some anecdotal evidence.

References

Airbnb (2022): "Shareholder Letter Q4 2021," .

Athey, S., J. C. Castillo, and B. Chandar (2021): "Service Quality on Online Platforms: Empirical Evidence about Driving Quality at Uber," *Working Paper*.

Aumann, R. J., M. Maschler, and R. E. Stearns (1995): *Repeated games with incomplete information*. MIT press.

Bergemann, D., and S. Morris (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57(1), 44–95.

Bergemann, D., and M. Ottaviani (2021): "Information markets and nonmarkets," in *Handbook of Industrial Organization*, vol. 4, pp. 593–672. Elsevier.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890.

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262.

Bertomeu, J., P. Ma, and I. Marinovic (2020): "How often do managers withhold information?," *The Accounting Review*, 95(4), 73–102.

Bertomeu, J., I. Marinovic, S. J. Terry, and F. Varas (2022): "The dynamics of concealment," *Journal of Financial Economics*, 143(1), 227–246.

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2017): "Julia: A fresh approach to numerical computing," *SIAM Review*, 59(1), 65–98.

Blackwell, D. (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, pp. 265–272.

Boleslavsky, R., and C. Cotton (2015): "Grading Standards and Education Quality," *American Economic Journal: Microeconomics*, 7(2), 248–79.

Boleslavsky, R., and K. Kim (2021): "Bayesian Persuasion and Moral Hazard," *Working Paper*.

Bundeskartellamt (2020): "Konsultationspapier zur Sektoruntersuchung Nutzerbewertungen," .

Calzolari, G., and A. Pavan (2006): "On the optimality of privacy in sequential contracting," *Journal of Economic Theory*, 130(1), 168–204.

Carnehl, C., M. Schaefer, A. Stenzel, and K. D. Tran (2021): "Value for Money and Selection: How Pricing Affects Airbnb Ratings," *Working Paper*.

Carnehl, C., A. Stenzel, and P. Schmidt (2021): "Pricing for the Stars–Dynamic Pricing in the Presence of Rating Systems," *Working Paper*.

Doval, L., and V. Skreta (2022): "Constrained Information Design," *Working Paper*.

Dranove, D., and G. Z. Jin (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48(4), 935–963.

Duffie, D., P. Dworczak, and H. Zhu (2017): "Benchmarks in Search Markets," *The Journal of Finance*, 72(5), 1983–2044.

Dworczak, P., and A. Kolotilin (2019): "The Persuasion Duality," *Working Paper*.

Dworczak, P., and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127(5), 1993–2048.

Dye, R. A. (1985): "Disclosure of nonproprietary information," *Journal of Accounting Research*, pp. 123–145.

Farronato, C., and A. Fradkin (2018): "The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb," *NBER Working Paper 24361*.

——— (2022): "The Welfare Effects of Peer Entry: The Case of Airbnb and the Accommodation Industry," *American Economic Review*, 112(6), 1782–1817.

Filippas, A., J. J. Horton, and J. M. Golden (2022): "Reputation inflation," *Marketing Science*.

Fradkin, A., E. Grewal, and D. Holtz (2021): "Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb," *Marketing Science*, 40(6), 1013–1029.

Fréchette, G. R., A. Lizzeri, and J. Perego (2022): "Rules and commitment in communication: An experimental analysis," *Econometrica*, 90(5), 2283–2318.

Gentzkow, M., and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review: Papers and Proceedings*, 106(5), 597–601.

GOFFE, W. L. (1996): "SIMANN: A global optimization algorithm using simulated annealing," *Studies in Nonlinear Dynamics & Econometrics*, 1(3).

GOFFE, W. L., G. D. FERRIER, AND J. ROGERS (1994): "Global optimization of statistical functions with simulated annealing," *Journal of Econometrics*, 60(1-2), 65–99.

HE, S., B. HOLLENBECK, AND D. PROSERPIO (2021): "The market for fake reviews," *Working Paper.*

HOPENHAYN, H., AND M. SAEEDI (2022a): "Optimal Coarse Ratings," *Working Paper.*

——— (2022b): "Optimal Information Disclosure and Market Outcomes," *Working Paper.*

HU, N., J. ZHANG, AND P. A. PAVLOU (2009): "Overcoming the J-shaped Distribution of Product Reviews," *Communications of the ACM*, 52(10), 144–147.

KAMENICA, E. (2019): "Bayesian Persuasion and Information Design," *Annual Review of Economics*, 11(1).

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101(6), 2590–2615.

KLEIN, T. J., C. LAMBERTZ, AND K. O. STAHL (2016): "Market Transparency, Adverse Selection, and Moral Hazard," *Journal of Political Economy*, 124(6), 1677–1713.

KOLOTILIN, A. (2018): "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13(2), 607–635.

KOLOTILIN, A., T. MYLOVANOV, AND A. ZAPECHELNYUK (2022): "Censorship as optimal persuasion," *Theoretical Economics*, 17(2), 561–585.

LIZZERI, A. (1999): "Information Revelation and Certification Intermediaries," *The RAND Journal of Economics*, 30(2), 214–231.

LUCA, M. (2016): "Designing Online Marketplaces: Trust and Reputation Mechanisms," *NBER Working Paper 22616.*

LUTZ, C., AND B. SAND (2019): "Highly disaggregated land unavailability," *Working Paper.*

Mayzlin, D., Y. Dover, and J. Chevalier (2014): "Promotional Reviews: An Empirical Investigation of Online Review Manipulation," *American Economic Review*, 104(8), 2421–55.

Ostrovsky, M., and M. Schwarz (2010): "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, 2(2), 34–63.

Rayo, L., and I. Segal (2010): "Optimal information disclosure," *Journal of Political Economy*, 118(5), 949–987.

Romanyuk, G., and A. Smolin (2019): "Cream Skimming and Information Design in Matching Markets," *American Economic Journal: Microeconomics*, 11(2), 250–76.

Rossi, M. (2021): "Competition and Reputation in a Congested Marketplace: Theory and Evidence from Airbnb," *Working Paper*.

Saeedi, M., and A. Shourideh (2020): "Optimal Rating Design," *arXiv preprint arXiv:2008.09529*.

Schaefer, M., and K. D. Tran (2020): "Airbnb, hotels, and localized competition," *DIW Berlin Discussion Paper 1889*.

Schoenmüller, V., O. Netzer, and F. Stahl (2018): "The extreme distribution of online reviews: Prevalence, drivers and implications," *Columbia Business School Research Paper*, (18-10).

Small, K. A., and H. S. Rosen (1981): "Applied welfare economics with discrete choice models," *Econometrica*, pp. 105–130.

Tadelis, S. (2016): "Reputation and feedback systems in online platform markets," *Annual Review of Economics*, 8, 321–340.

Teubner, T., F. Hawlitschek, and D. Dann (2017): "Price determinants on Airbnb: How reputation pays off in the sharing economy," *Journal of Self-Governance and Management Economics*, 5(4), 53–80.

Train, K. (2015): "Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples," *Journal of Choice Modelling*, 16, 15–22.

Vatter, B. (2021): "Quality disclosure and regulation: Scoring design in medicare advantage," *Working Paper*.

Vellodi, N. (2021): "Ratings design and barriers to entry," *Working Paper*.

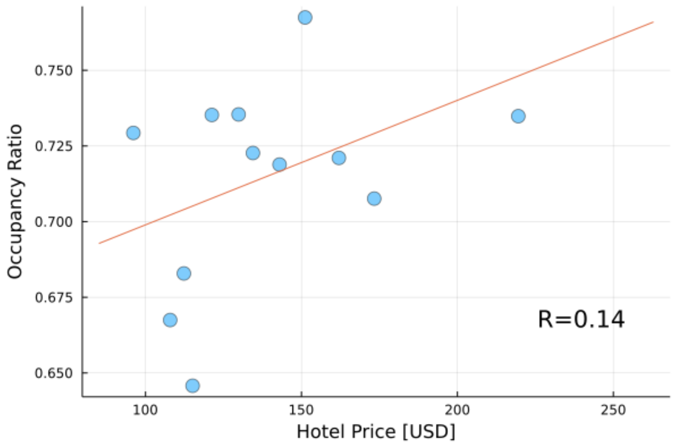
Wilson, R. (1989): "Efficient and Competitive Rationing," *Econometrica*, 57(1), 1–40.

Zervas, G., D. Proserpio, and J. W. Byers (2021): "A first look at online reputation on Airbnb, where every stay is above average," *Marketing Letters*, 32(1), 1–16.

Appendix

A. Additional Figures and Tables

Figure A.1: Cross-Market Correlation of Hotel Prices and *occ_rat*



Notes: The graph shows binned scatter plots (blue circles), a linear fit line (red), and the Pearson correlation coefficient ('R') of the occupancy ratio between non-5-star and 5-star listings (occ_rat) with respect to the average hotel price across 112 markets on the horizontal axes.

Table A.1: Geographical Markets

urban (u)		non-urban (nu)
Atlanta, GA	Albuquerque, NM	Arizona Area
Boston, MA	Asheville, NC	Bend/Redmond, OR
Chicago, IL	Austin, TX	Colorado Area
Dallas, TX	Baltimore, MD	Corpus Christi, TX
Denver, CO	Charleston, SC	Daytona Beach, FL
Houston, TX	Charlotte, NC-SC	Florida Panhandle
Los Angeles/Long Beach, CA	Indianapolis, IN	Fort Lauderdale, FL
Miami/Hialeah, FL	Jersey City/Secaucus, NJ	Galveston/Texas City, TX
Nashville, TN	Louisville, KY-IN	Gatlinburg/Pigeon Forge, TN
New York, NY	Memphis, TN-AR-MS	Hilton Head/Beaufort, SC
Orlando, FL	Minneapolis/St Paul, MN-WI	Merced/Central CA
Philadelphia, PA-NJ	New Orleans, LA	Mobile, AL
Phoenix, AZ	Oakland, CA	Palm Springs, CA
San Antonio, TX	Pittsburgh, PA	Salt Lake City/Ogden, UT
San Diego, CA	Portland, OR	Santa Fe, NM
San Francisco/San Mateo, CA	San Jose/Santa Cruz, CA	Sarasota/Beaches, FL
Seattle, WA	St Louis, MO-IL	Savannah, GA
Washington, DC-MD-VA	Tucson, AZ	Tampa/St Petersburg, FL
		Texas West
		Utah Area

Notes: Geographical units are industry-standard STR markets/sub-markets. Urban/non-urban classification done by author.

Table A.2: First Stages

	(1)	(2)		(3)	(4)
	Airbnb			**Hotels**	
	Price	log within mkt share		Price	log within mkt share
rent	1.127***	0.0254***	share unavailable land	0.781***	-0.00101**
	(0.136)	(0.00570)		(0.268)	(0.000454)
share hotel rooms	6.389	1.334**	share Airbnb entire home	-49.39	-0.162**
$\geq$ upper-midscale	(28.01)	(0.545)		(46.92)	(0.0790)
Observations	112	112	Observations	112	112
R-squared	0.753	0.621	R-squared	0.413	0.353
Controls	yes	yes	Controls	yes	yes
Mean dep. var.	115.4	-2.488	Mean dep. var.	139.1	-0.108
F-Stat.	36.51	22.52	F-Stat.	4.256	8.399

Notes: Stadard errrors in parentheses clustered at geo-markets; *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Parameter Estimates *without Demand Instruments*

		Parameter Estimate	Standard Error
Demand Coefficients	Price (α) [4,5]	0.01349	0.00347
	FE urban destination (γ_u) [4,5]	0.55	0.47
	FE non-urban destination (γ_{nu}) [4,5]	0.51	0.53
	Log OSM tourist POIs (ψ) [4,5]	0.32	0.10
	Share 'private room' (β_1^A) [4]	-7.84	1.01
	Share 'entire home' studio (β_2^A) [4]	-0.79	1.55
	Share 'entire home' 2 bedroom (β_3^A) [4]	-5.71	1.01
	Avg. bookig.com rating (β_1^H) [5]	1.54	0.95
	Share $\geq$ Upper Midscale (β_2^H) [5]	0.82	0.35
	Nesting parameter (ζ) [4,5]	0.227	0.598
Unobs. Airbnb Quality	Mean (ν)	-1.50	0.49
	Std. dev. (σ)	1.27	0.30
Effort Costs	Constant (κ_0) [2]	56.92	2.69
	Wage coefficient (κ_1) [2]	0.13	1.79
Listing Costs	Constant (η_0) [3]	35.40	0.70
	Rent coefficient (η_1) [3]	-0.01	0.08

Notes: GMM estimates and standard errors of the model parameters without instrumenting for prices and within market shares in the demand equations. The numbers in brackets refer to the equations in the main text the parameters are defined in. Standard errors are calculated based on the usual analytical formula for GMM standard errors, based on numerical derivatives of the objective function at the optimum.

Table A.4: Average Demand Elasticities *without Demand Instruments*

	Airbnb	Hotel	Outside
Airbnb	-1.92	0.88	-
Hotel	0.09	-1.54	-
Outside	0.04	0.39	-

Notes: Nested logit demand elasticities for non-compression nights (i.e., nights in which customers have the choice between all three options), based on the estimation without instrumenting for prices and within market shares in the demand equations (Table A.3), averaged across 112 markets.

B. Omitted Derivations

In the following, market subscripts are omitted to improve readability.

B.1. Rearrangement of the Host FOC

In this section, I derive the expression for the representative host's first-order condition stated as equation 11 in the main text. I follow the same steps that Boleslavsky and Kim (2021) perform for the case with a discrete state space.

Given information structure π and conjectured effort level $\hat{e}$ and cutoff quality $\underline{\hat{\omega}}$, by Bayes' rule customers' posterior density for any $s \in \mathcal{S}$ is given by

$$\mu^{Bayes}(\omega|s, \pi, \hat{e}, \underline{\hat{\omega}}) = \frac{\pi(s|\omega) f_{\underline{\hat{\omega}}}(\omega; \hat{e})}{\int_{\underline{\hat{\omega}}}^{\infty} \pi(s|\omega') f_{\underline{\hat{\omega}}}(\omega'; \hat{e}) d\omega'}. \tag{B.1}$$

This defines customers' Bayes-plausible belief structure. The equilibrium belief structure M needs to be Bayes-plausible given equilibrium effort e, equilibrium cutoff quality $\underline{\omega}$, and equilibrium information structure π. Moreover, the conjectured effort level and cutoff quality must coincide with the hosts' optimal choices. Hence, we must have $e = \hat{e}$, $\underline{\omega} = \underline{\hat{\omega}}$, and $\mu_s(\omega) = \mu^{Bayes}(\omega|s, \pi, e, \underline{\omega})$. Moreover, note that the denominator of the right-hand side of equation B.1 describes the unconditional density of the Bayesian update μ_s, $\tilde{g}(\mu_s)$. Hence, equation B.1 can be rewritten as

$$\pi(s|\omega) = \frac{\mu_s(\omega) \tilde{g}(\mu_s)}{f_{\underline{\omega}}(\omega; e)}. \tag{B.2}$$

The first order condition of Π^{RH} as defined in the first line of equation 9 in the main text with respect to e reads as

$$(1-\tau)(p^H + p) \int_{\underline{\omega}}^{\infty} \int_{\mathcal{S}} \pi(s|\omega) \rho^A \left(\mathbb{E}_{\mu_s}[\omega]\right) ds \frac{\partial f(\omega; e)}{\partial e} d\omega = T f(\underline{\omega}; e) + 2ke. \tag{B.3}$$

Using B.2 and switching the order of integration, B.3 can be restated in terms of posterior beliefs only instead of signals as

$$(1-\tau)(p^H + p) \int_{\mathcal{S}} \rho^A \left(\mathbb{E}_{\mu_s}[\omega]\right) \mathbb{E}_{\mu_s} \left[\frac{\partial f(\omega; e)/\partial e}{f_{\underline{\omega}}(\omega; e)} \right] d\tilde{G}(\mu_s) = T f(\underline{\omega}; e) + 2ke,$$

which corresponds to equation 11 in the main matter.

B.2. Derivation of the Numerical Solution Algorithm

By Theorem 3 and Theorem 4 of Dworczak and Kolotilin (2019) for a given price and target effort level, the platform's optimization problem 13 of finding the profit-maximizing distribution of posterior means G is equivalent to the dual problem of finding an optimal price function $\varphi \in C([\underline{\omega}, \infty))$ for posterior means and a non-negative scalar λ that solve

$$\begin{aligned} \min_{\varphi,\lambda} & \int_{\underline{\omega}}^{\infty} \varphi(m) dF_{\underline{\omega}}(m;e) - \lambda \frac{\sigma^2[Tf(\underline{\omega};e) + 2ke]}{(1-\tau)(p^H + p)[1 - F(\underline{\omega};e)]} \\ \text{s.t. } & \varphi(m) - \lambda\rho^A(m)(m - \nu - e) \geq \rho^A(m) \quad \forall m \in [\underline{\omega}, \infty) \end{aligned} \tag{B.4}$$

and in the optimum, we have

$$\int_{\underline{\omega}}^{\infty} \varphi(m) dF_{\underline{\omega}}(m;e) - \lambda \frac{\sigma^2[Tf(\underline{\omega};e) + 2ke]}{(1-\tau)(p^H + p)[1 - F(\underline{\omega};e)]} = \int_{\underline{\omega}}^{\infty} \rho^A(m) dG(m).$$

If I fix λ, I can rewrite the dual problem as

$$\begin{aligned} \min_{\varphi} & \int_{\underline{\omega}}^{\infty} \varphi(m) dF_{\underline{\omega}}(m;e) \\ \text{s.t. } & \varphi(m) \geq [1 + \lambda(m - \nu - e)]\, \rho^A(m) \quad \forall m \in [\underline{\omega}, \infty). \end{aligned}$$

This is (see Theorem 3 by Dworczak and Kolotilin (2019)) the dual problem of an unconstrained Bayesian persuasion problem with the payoff function

$$\tilde{v}(m) = [1 + \lambda(m - \nu - e)]\, \rho^A(m),$$

which depends on posterior beliefs only through posterior means m.

Hence, for any given λ, if $\tilde{v}(m)$ is S-shaped, the optimal information structure is upper-censoring, and the threshold quality ω^* can be found by solving equation 1 (replacing $\rho^A(m)$ by $\tilde{v}(m)$). While the conditional choice probability functions resulting from a logit or nested logit demand system would be S-shaped, $\rho^A(m)$ as conditional choice probability resulting from a mixture of a three-alternative nested logit system (non-compression nights) and a two-alternative logit system (compression nights) is in general not S-shaped. However, as regular checks during the optimization algorithm show, the S-shape is preserved for all empirically relevant parameter values. The penalized objective function $\tilde{v}(m)$ perturbs $\rho^A(m)$ in a way such that the S-shape may be violated and $\tilde{v}(m)$ can be concave for very small m and convex for very large m if λ is large. In all empirically relevant cases, the initial concave portion lies below the participation cutoff $\underline{\omega}$ and is therefore irrelevant.

Moreover, in all empirically relevant cases, the convexity at the top is not strong enough to provoke an additional interval of full revelation. Hence, the upper-censoring information structure remains optimal and can be determined as for the case with an S-shaped objective function.

In summary, the numerical solution algorithm comprises three nested optimization loops. In the outer loop, I maximize Π^A over p and the targeted effort level e. In the middle loop, for fixed p and e, I determine the λ that minimizes B.4. In the inner loop, I solve for the optimal upper-censoring information structure in the unconstrained problem for fixed p, e, and λ by solving equation 1 from the main text with $\tilde{v}(m)$ as objective function.

B.3. Calculation of Outcome Measures for Counterfactual Experiments

This appendix section provides some more details on calculating the outcome measures for the counterfactual experiments presented in Section 7 of the main text.

In the counterfactual scenarios with fully revealing ratings (columns 3 and 5 of Table 5), the distribution of posterior means corresponds to the prior distribution. So, market shares are calculated by integrating the respective choice probability function conditional on posterior means with respect to the prior distribution. For Airbnb, this yields

$$\rho^A = \int_{\underline{\omega}}^{\infty} \rho^A(m) dF(m; e),$$

and the hotel market share is obtained in the same way.

In the counterfactual scenario with no ratings and therefore no participation restriction (column 4 of Table 5), market shares are calculated by evaluating the conditional choice probability functions at the mean ν of the prior distribution without any additional host effort. That is, for Airbnb

$$\rho^A = \rho^A(\nu),$$

and accordingly for hotels.

When keeping the baseline levels of Airbnb price and participation cutoff fixed (column 6 of Table 5), the market shares without ratings are calculated by evaluating the conditional choice probability functions at the mean of the truncated prior distribution without any additional host effort, $\mathbb{E}_{F(\cdot;0)}[\omega|\omega > \underline{\omega}]$, and accounting for the truncation by multiplying with $1 - F(\underline{\omega}; 0)$. For Airbnb, we get

$$\rho^A = (1 - F(\underline{\omega}; 0))\, \rho^A \left(\mathbb{E}_{F(\cdot;0)}[\omega | \omega > \underline{\omega}] \right),$$

and accordingly for hotels.

In the case with binary signals and fixed prices and host participation (Appendix C), we get the Airbnb market share

$$\begin{aligned} \rho^A = &\left(F(\omega^{bin}; e) - F(\underline{\omega}; e) \right) \rho^A \left(\mathbb{E}_{F(\cdot;e)}[\omega | \underline{\omega} < \omega < \omega^{bin}] \right) \\ &+ \left(1 - F(\omega^{bin}; e) \right) \rho^A \left(\mathbb{E}_{F(\cdot;e)}[\omega | \omega > \omega^{bin}] \right), \end{aligned}$$

where ω^{bin} is the threshold value separating the low and high binary ratings. The hotel market share is calculated in the same fashion.

The expected profit functions for Airbnb and the representative host in the different counterfactual experiments are obtained by replacing in equations 8 and 9 the upper-censoring market shares with the market shares in the respective counterfactual scenario calculated according to the formulas above.

For the specified demand system accounting for hotels' capacity constraints and different choice sets for the marginal Airbnb customer in compression and non-compression nights, in the baseline case of upper-censoring ratings, the expected indirect utilities for the calculation of the Airbnb customer welfare measure in equation 14 are given by

$$\begin{aligned} \mathbb{E}[\max\{U^A, U^H, U^0\}] = \; & scn\; \bar{\rho}^H \int_{\underline{\omega}}^{\infty} \log\left(1 + \exp\{\delta^A + m\}\right) dG(m) \\ & + (1 - scn) \Bigg[F(\underline{\omega}; e) \log\left(1 + exp\{\delta^H\}\right) \\ & + \int_{\underline{\omega}}^{\infty} \log\left(1 + \left(\exp\left\{\frac{\delta^H}{1-\zeta}\right\} + \exp\left\{\frac{\delta^A + m}{1-\zeta}\right\}\right)^{1-\zeta}\right) dG(m) \Bigg] \end{aligned} \tag{B.5}$$

and

$$\mathbb{E}[\max\{U^H, U^0\}] = (1 - scn) \log\left(1 + exp\{\delta^H\}\right).$$

In the different counterfactual scenarios, $E[\max\{U^A, U^H, U^0\}]$ is calculated accordingly by integrating in equation B.5 with respect to the perspective induced distribution of posterior means, as for the calculation of the market shares resented at the beginning of this section.

C. Binary Ratings

This section considers two additional counterfactual exercises with platform-optimal and consumer-optimal binary ratings. That is, unlike before, I restrict the signal space to two discrete signals (e.g., thumb up/thumb down), and for each market, I find the cutoff quality ω^{bin} that maximizes expected platform profits or customers' CV for Airbnb, respectively. The 'high' rating induces a posterior believed quality of $\mathbb{E}_{F(\cdot;e)}[\omega|\omega > \omega^{bin}]$ while the 'low' rating induces a posterior believed quality of $\mathbb{E}_{F(\cdot;e)}[\omega|\underline{\omega} < \omega < \omega^{bin}]$.

To isolate the role of ratings in identifying qualities of actually listed properties, in both scenarios, I keep Airbnb prices and host participation (i.e., $\underline{\omega}$) as in the baseline with platform optimal upper-censoring ratings, which by assumption correspond to the status quo. The only choice variable is ω^{bin}, and the corresponding host effort level is found by solving the representative host's first-order condition (equation 12 in the main text). Market shares and profits with binary ratings are calculated as explained in Appendix B.3.

Table C.1, which is organized as Table 5 in the main text, summarizes the market outcomes and welfare measures under these counterfactual scenarios. Column 1 shows the baseline levels under platform-optimal upper-censoring ratings. Column 2 shows the outcomes for platform-optimal binary ratings (i.e., ω^{bin} is chosen to maximize expected platform profits), and column 3 shows the outcomes for customer-optimal binary ratings (i.e., ω^{bin} is chosen to maximize customers' CV for Airbnb). As additional points of reference, column 4 repeats column 3 from Table 5 with no ratings and fixed prices and host participation, and column 5 adds the case with fully revealing ratings and fixed prices and host participation.

While with platform-optimal upper-censoring ratings, on average, 72% of listings have a 5-star rating, with platform-optimal binary ratings, we have on average 74% listings with the 'high' rating and with customer-optimal binary ratings, we have on average only 27% listings with the 'high' rating. Customers benefit much more from being able to identify the very best listings among all the acceptable ones than from being able to identify the worst qualities among all the acceptable ones (recall that host participation is fixed in this counterfactual).

Overall, the platform-optimal binary ratings (column 2) are a very close approximation of the upper-censoring ratings. The welfare consequences work mainly through less induced host effort but are very small on the whole. Both Airbnb and customers only benefit marginally from knowing the exact qualities of the bottom listings instead of only knowing that they are below the cutoff value. However, keep in mind that taking the model at face value, binary ratings alone cannot exclude the worst qualities from participation as they do not fully reveal the quality around the threshold $\underline{\omega}$. In this sense, all the scenarios with fixed host participation are not inherently stable outcomes but only hypothetical scenarios to isolate the purely informative role of ratings in identifying different qualities among those actually listed. A counterfactual simulation for a stable outcome with binary ratings would require an additional modeling assumption on how participation adjusts if ratings are not fully revealing around the participation threshold $\underline{\omega}$.

Remarkably, customer-optimal binary ratings (column 3), despite only using two signals instead of 9 or a continuum, can improve the ex-ante Airbnb value to customers by 1.79% compared to the status quo and can thereby deliver more than half the benefit of fully-revealing continuous signals (compare to column 5). This shows that the discrete nature of the rating scale or the number of available signals on a 5-star scale are not the primary factors limiting information transmission to benefit customers. It is instead the design by the platform which determines the cutoff values in a strategic way that increases their profits but harms the average customer.

The results for the scenario with fully revealing ratings but with Airbnb prices and participation cutoff $\underline{\omega}$ fixed as in the baseline (column 5) demonstrate again the substitutability between higher prices and less information. If the platform can re-adjust prices, it can get back parts of the lost profits from fully revealing ratings by increasing the price (recall column 4 of Table 5 in the main text). If Airbnb is not allowed to adjust the price, all the welfare consequences are more pronounced.

Table C.1: Counterfactuals with Binary Ratings

	(1) **Baseline** (Platform-optimal ratings)	(2) **Platform-Optimal Binary Ratings** (Price & participation fixed as in baseline)	(3) **Customer-Optimal Binary Ratings** (Price & participation fixed as in baseline)	(4) **No Ratings** (Price & participation fixed as in baseline)	(5) **Fully Revealing Ratings** (Price & participation fixed as in baseline)
Airbnb Price [US$]	117.81	117.81	117.81	117.81	117.81
		0%	0%	0%	0%
Effort [US$]	0.95	0.94	0.98	0	1.01
		-0.61%	+3.70%	-100%	+7.13%
Airbnb Market Share	0.0294	0.0294	0.0292	0.0284	0.0291
		-0.01%	-0.52%	-3.27%	-0.89%
Hotel Market Share	0.2147	0.2147	0.2148	0.2150	0.2148
		-0.001%	+0.03%	+0.11%	+0.05%
Platform Profit per	0.57	0.57	0.56	0.55	0.56
***Potential* Listing [US$]**		-0.01%	-0.52%	-3.31%	-0.89%
Platform Profit per	9.28	9.28	9.23	9.27	9.20
***Actual* Listing [US$]**		+0.003%	-0.47%	-0.12%	-0.83%
Host Profit per	0.91	0.91	0.89	0.89	0.88
***Potential* Listing [US$]**		+0.04%	-2.13%	-2.07%	-3.76%
Host Profit per	13.42	13.42	13.18	13.51	12.99
***Actual* Listing [US$]**		+0.02%	-1.78%	+0.69%	-3.18%
Ex-ante Customer CV	2.76	2.76	2.81	2.61	2.86
for Airbnb [US$]		-0.21%	+1.79%	-5.47%	+3.51%
Airbnb Customer CV	92.60	92.47	94.40	91.04	96.02
per Booked Night [US$]		-0.14%	+1.94%	-1.68%	+3.69%

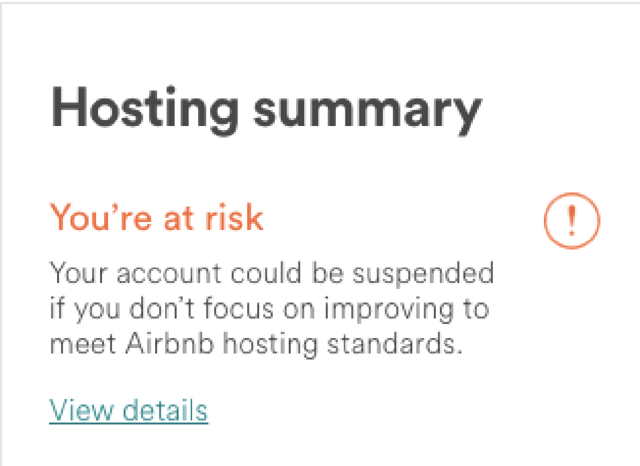
Notes: Averages across 112 markets.

D. Additional Anecdotal Evidence

D.1. Airbnb

Platform Only Accepts Five Stars

Figure D.1: Airbnb pop-up message for a host with insufficient average rating



Source: https://medium.com/@campbellandia/how-to-avoid-the-dreaded-4-star-review-a-guide-for-airbnb-hosts-cdf482d083fe (accessed on September 24, 2022).

Rating Instructions by Airbnb Hosts

Figure D.2: Rating Instructions by Airbnb Hosts – II

Please be aware ...
of the judgement the stars may bear

When rating us after your stay, please consider that Airbnb's rating is totally different than the classic hotel star rating system: You should know that Airbnb can close down any listing that drops below an average of 4.3 stars at once. So the least we can say is that the system is rather confusing. To give you a better understanding, we designed the following (cheeky) guide to the rating system:

★★★★★	Perfection does not exist, but we were happy here.
★★★★	Several issues need to be improved.
★★★	Major problems that should be fixed immediately.
★★	Close it down at once.
★	Put these hosts into jail.

We always strive to give our guests a 5 star experience. Please let us know any problem you may have during your stay and we will do everything within our power to solve it.

Source: https://airhostsforum.com/t/review-star-system-need-to-educate-guests/28227/3 (accessed on September 24, 2022).

Figure D.3: Rating Instructions by Airbnb Hosts – III

airbnb

Superhost

IMPORTANT NOTE ON AIRBNB RATINGS – PLEASE READ!

As a Superhost, I do everything I can to create a positive experience and ensure that my cottage meets your expectations based on my listing description, location and price. Good reviews enable me to continue offering my home as a safe, comfortable and economical place for visitors to stay.

Airbnb ratings are not like hotel ratings – please don't compare my humble accommodation to the Ritz! A 5-star rating means that you were happy with your stay *based on the price you paid.* Airbnb considers a score of less than 5 as a hosting 'fail'– if my average rating drops below 4.8, I will lose my Superhost status, and if it drops below 4.7, I am threatened with being delisted altogether! (This is no joke). Hosts also lose their Superhost status if their review rate drops below 50% - I review all my guests (with almost all receiving 5 stars) and I would very much appreciate your review in return!

The 'Overall Experience' star rating is based on the question, "How did your stay at **Rob's** place compare to your expectations?"

Possible guest answers	Overall rating given to host	How Airbnb rates the host
Much better than I expected	5 stars ★★★★★	Acceptable
A bit better than I expected	4 stars ★★★★	Not good enough
About the same as I expected	3 stars ★★★	Bad
A bit worse than I expected	2 stars ★★	Terrible **(suspended)**
Much worse than I expected	1 star ★	Totally unacceptable **(delisted)**

Of course, if you are disappointed or dissatisfied with any aspect of your stay, please feel free to communicate this to me so I have a chance to address the issue and ensure your stay is a 5-star experience!

"Please feel at home & if there is anything you would like that is not at hand, just ask"!

Rob & Ade

Source: https://community.withairbnb.com/t5/Hosting/Explaining-5-star-ratings-to-guests/td-p/882631 (accessed on September 24, 2022). Picture of hosts blurred by the author.

D.2. Uber

Explicit Instructions

In 2019, Uber wrote explicitely on their customer help pages:

> *Most riders provide a 5-Star rating unless there was a specific issue with the trip. If we see a 1-Star rating, it typically means that there is a serious problem with a driver.*[54]

Platform Only Accept Five Stars

Uber writes in their *Community Guidelines*:

> *There is a minimum average rating in each city. [...] Drivers [...] that don't meet the minimum average rating for their city may lose access to all or part of the Uber Marketplace Platform. If your rating is approaching this limit, we will let you know and may share information that may help you improve your rating.*[55]

A Frequent Customer's View

> *[...] in my experience, the lack of ratings differentiation makes it difficult to distinguish between exceptional and marginally acceptable service. I take an embarrassing number of Uber and Lyft rides each week. I've ridden with 4.7 star drivers who wear gloves and open passenger doors and 4.7 star drivers who couldn't pass a road test. It may seem a petty distinction, but in some cases, inflated ratings could threaten customer safety.*

Kat Kane, wired.com, 19.03.2015[56]

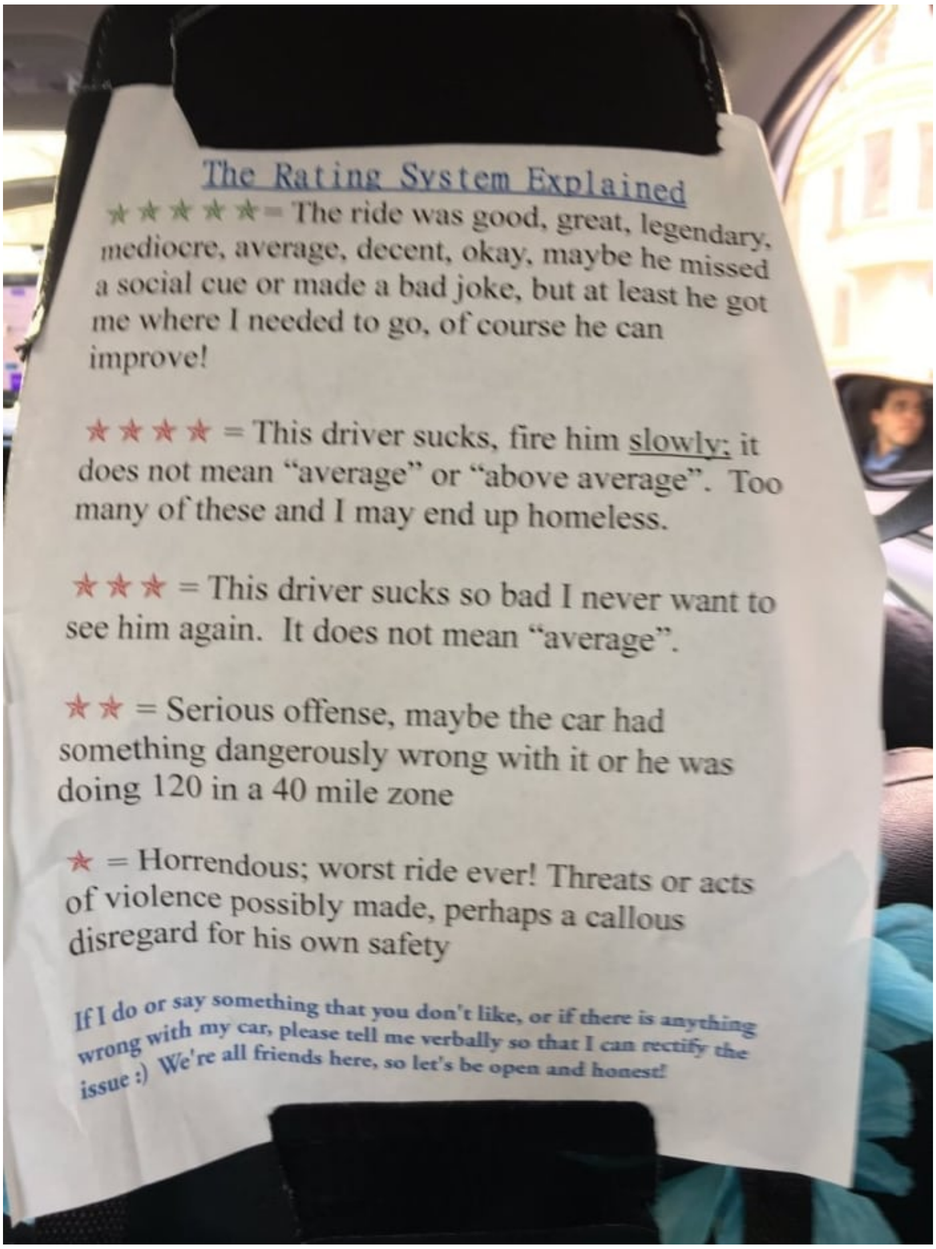
[54] Source: https://help.uber.com/riders/article/rating-a-driver?nodeId=478d7463-99cb-48ff-a81f-0ab227a1e267 (accessed on June 17, 2019). In the time following, the exact wording has been modified several times, and by now these explicit rating instructions have been removed completely; see https://web.archive.org/web/20200801000000*/https://help.uber.com/riders/article/rating-a-driver?nodeId=478d7463-99cb-48ff-a81f-0ab227a1e267.

[55] Source: https://www.uber.com/legal/community-guidelines/us-en/ (accessed on September 24, 2022).

[56] Source: https://www.wired.com/2015/03/bogus-uber-reviews/ (accessed on September 24, 2022).

Rating Instructions by Uber Drivers

Figure D.4: Rating Instructions by Uber Drivers



Source: Caroline O'Donovan, BuzzFeed News, https://www.buzzfeednews.com/article/carolineodonovan/the-fault-in-five-stars (accessed on September 24, 2022).