The Cost of the Cold-Start Problem on Airbnb

Florian Dendorfer and Regina Seibel*

October 23, 2024

Abstract

In digital markets with peer-to-peer reviews, new products encounter the so-called "cold-start" problem: Little-known products are bought too rarely and remain little known. In this paper, we examine the inefficiency associated with social learning on Airbnb, including its implications for hosts' price, entry and exit decisions. We estimate a dynamic structural model of demand and supply for Airbnbs in Manhattan, New York. We find that addressing the cold-start problem by lowering the price of new listings relative to incumbent ones leads to a welfare increase exceeding 8% of total host revenue.

Keywords: Cold-start problem, digital platforms, experimentation, market dynamics, product reviews, social learning.

JEL: L11, L15, L83, L86, L88, D83

1 Introduction

A key characteristic of digital platforms is their virtually unlimited shelf space, which provides consumers with an extensive range of products. Amid this abundance, peer-to-peer reviews have become crucial, allowing consumers to learn from the experiences of previous buyers and improve their purchasing decisions. However, reviews are likely to be misallocated. Consumers often hesitate to purchase unreviewed products, which might be of poor quality, preferring reviewed alternatives. At the same time, the marginal review of a little-known product provides more information about its quality than that of a well-established product. Therefore, its social value is relatively larger, but this is not reflected in consumer purchase decisions. All else equal, the lack of exploration results in the "cold-start" problem (Che and Hörner, 2018; Kremer, Mansour and Perry, 2014), where new products are discovered at an insufficient rate initially. Indeed, many platforms, such as Airbnb and Amazon, recognize the challenges and significance

*Dendorfer: Department of Economics, University of Toronto. Seibel: Rotman School of Management, University of Toronto. Email: `florian.dendorfer@utoronto.ca`, and `regina.seibel@rotman.utoronto.ca`. We thank Heski Bar-Isaac, Stefan Bühler, El Hadi Caoui, Gregory Crawford, Florian Ederer, Luise Eisfeld, Brett Hollenbeck, Mitsuru Igami, Armin Schmutzler, and conference participants at CRESSE 2023, IIOC 2023, MaCCI 2023, the Swiss IO Day 2023 and the 6th Doctoral Workshop on the Economics of Digitization for helpful comments.

of accumulating initial reviews and seem to believe that addressing the cold-start problem is beneficial for the marketplace they operate.[1]

However, whether increasing the speed of learning about new products enhances overall welfare becomes less clear when taking the endogenous decisions of sellers into account.[2] As sellers set prices, and decide whether to enter or exit the market, they consider the impact of reviews on their lifetime profit. For instance, new sellers might lower their prices to increase their chances of receiving reviews and gaining a competitive edge, a strategy akin to learning-by-doing (Cabral and Riordan, 1994, 1997). Any intervention aimed at addressing the cold-start problem could inadvertently distort seller profits, impacting their decision to enter or exit the market. Consequently, while some seller decisions can alleviate the welfare implications of the cold-start problem, others may exacerbate it, complicating any general statements about its existence and impact.

In this paper, we examine the cold-start problem empirically and analyze its impact on consumers, sellers, and overall welfare. Specifically, we estimate a structural model of the Airbnb market in Manhattan, New York, that includes dynamic supply-side decisions. Listing quality is initially unknown, and peer-to-peer reviews result from past purchases. Using this model, we explore counterfactuals which induce sellers to change their prices, shifting demand between sellers and alleviating the welfare loss from the cold-start problem, i.e., from insufficient social learning. We find that redirecting demand from established listings to new ones, thereby increasing the speed of social learning, increases total welfare. However, the responses of sellers in term of entry and exit play a major role in mediating the welfare change as they have a first-order effect on listing variety. To our knowledge, we provide the first structural empirical analysis of the existence and extent of the cold-start problem.

To illustrate the cold-start problem, we begin by analyzing a simple two-period model of social learning involving only two firms: an incumbent and an entrant, with the quality of the entrant's product being a priori unknown. From a social surplus perspective, the entrant's price in the first period should be lower than the incumbent's price. This is not always true in the Nash equilibrium: entrants charge a higher price if they are expected to be of sufficiently high quality. Dynamic pricing considerations increase the price difference between incumbents and entrants, but the price difference remains insufficiently low if the entrant's expected quality is not much lower than the incumbent's quality. We conclude our simple model by discussing the implications of incorporating an outside option, as well as seller entry and exit. In particular, entry and exit incentives introduce concerns of market participation, which are related by the cold-start problem in a non-trivial way. Our analysis suggests that the existence of the cold-start problem may not only depend on the prior beliefs about product quality, but is also closely tied

[1] Airbnb recommends hosts that they offer a 20% discount to their very first guests, see https://www.airbnb.ca/resources/hosting-homes/a/how-to-set-a-pricing-strategy-15. Amazon allows sellers to offer their first products for free in exchange for guaranteed reviews from trusted customers on the platform, see https://sell.amazon.com/tools/vine.

[2] Bergemann and Välimäki (1997, 2000) incorporate endogenous pricing and find that the speed of social learning is too high, not too low. Hagiu and Wright (2020) find that social learning may be too high or too low from a platform perspective if sellers have market power or the platform puts a positive weight on buyer surplus. Vellodi (2022) demonstrates that allowing for endogenous pricing fully alleviates the cold-start problem in his model with endogenous entry.

to entry and exit dynamics.

We then introduce our empirical model of the Airbnb market. Airbnb guests make individually optimal booking choices among the available listings based on their expectation about a listing's quality. When deciding which listing to book, guests do not take into account the impact of the review that they might write on future guests. Guests hold prior beliefs that they update according to Bayes' rule based on the number of reviews and the rating they observe. On the supply side, there is a pool of hosts, some of whom are inactive. Active hosts set rental prices while considering the potential effect on future demand through the accumulation of reviews. Hosts also make entry and exit decisions; active hosts decide whether to remain active or exit the market and save the operating cost, while inactive hosts can enter at the beginning of each period at an entry cost. Given the large number of Airbnb hosts, we approximate the symmetric Markov equilibrium of this game using the oblivious equilibrium concept introduced by Weintraub, Benkard and Van Roy (2008), where hosts assume that the distribution of competitors remains fixed and equal to the steady state.

We estimate the demand and supply sides of our model separately. To recover the demand parameters, we employ the Generalized Method of Moments. Our estimates underscore the significant impact that ratings have on both guest and host behavior. On average, a good review increases a listing's occupancy rate by 2.6%. For unreviewed listings, this effect is more than twice as large. We estimate a mean quality corresponding to 4.15 stars (on a one-to-five-star scale) in the listing population, whereas the average rating in the data is 4.51 stars. This discrepancy arises because only high-quality listings tend to survive in the market.

We obtain the supply parameters by using a nested maximum likelihood routine to match the observed distribution of listing types and reviews. Our estimates indicate that hosts face substantial entry costs, averaging about half of their lifetime profit. Based on our estimates of hosts' operating costs, we calculate an average profit margin of 24.5%. To evaluate model fit, we use our model and the parameter estimates to simulate data and compare it to our sample. The rating distribution, rental prices, and occupancy rates in the simulated data align well with the actual data, whereas the model predicts too large exit rates. We demonstrate that reviews influence host and consumer behavior in a manner broadly consistent with the observed data.

In the counterfactual analysis section of this paper, we induce price adjustments which maximize long-run equilibrium welfare. To reduce the dimensionality of the problem, we group listings based on their number of reviews and employ two-part tariffs consisting of subsidies or taxes and lump-sum transfers. Subsidies and taxes are uniform within each group but may vary across groups. Hosts of taxed listings are compensated for the tax burden, and hosts of subsidized listings are charged for the subsidy cost through lump-sum transfers. With our approach, we effectively compel hosts to set a price that is different from their individually optimal price. To ensure that the welfare effects we observe are due to changes in the appeal of entrants relative to incumbents, rather than a change in the appeal of any Airbnb listings relative to the outside good, we impose the condition that the average tax-subsidy must be zero. In a second counterfactual, inspired by Airbnb's attempt to influence prices of brand-new listings, we directly search over the socially optimal prices of listings with no reviews.

In the first counterfactual, we find that the welfare-maximal tax-subsidy scheme causes listings with at most five reviews to decrease their price, while listings with more than five reviews increase their price. The largest price changes are observed for listings with at most one review (-16.6%) and listings with more than 15 reviews (+9.7%). This increase in the price gap between relatively established and new listings diverts demand towards the newer listings. Hence our policy increases the speed of social learning. However, socially optimal prices are also influenced by consideration for host exit and entry decisions. Active incumbent listings are more valuable to consumers than entrant listings, and our policy reconciles motives of increasing the speed of learning and avoiding market exit by incumbents. Widening the price gap benefits incumbent hosts because it allows them to charge higher prices. In the new equilibrium, we observe 18% fewer listings with at most one review but over 50% more listings with more than 15 reviews. This results in a welfare increase by almost $340,000 per month, an increase that is mostly driven by an increase in consumer surplus. Not only consumers, but also Airbnb benefit from the policy, with an 8% increase in its revenue from Airbnb fees. At the same time, hosts are hurt by the policy overall: they lose more than 8% of their surplus.

While in our first counterfactual, the motive of ensuring market participation of socially more valuable hosts aligns with the motive of alleviating the cold-start problem, this is not the case in our second counterfactual. There, we focus on listings with no reviews and induce only them to set the socially optimal prices. As a result, we observe an interesting trade-off. Similar to our first counterfactual, the price for these new listings is lower than in the status quo. Since new listings are forced to reduce their prices, their profits drop and their incentive to enter decreases. At the same time, there is no compensating decrease in exit from other hosts. Overall, this policy leads to rather mild welfare increase, as reduced entry incentives and the subsequent loss in variety counteract the gains from social learning.

Our analysis indicates that the cold-start problem may lead to a significant welfare loss on Airbnb and potentially other markets with similar characteristics, though the extent of the cold-start problem generally depends on entry and exit dynamics, as well as the prior expectation of entrants' quality relative to incumbents'. For the market we analyze, we find that products with few reviews are priced too high relative to the ones with many reviews, resulting in too little demand and slow learning. Consequently, consumers have inefficiently little information about product quality.

The paper is organized as follows. In Section 2, we discuss the relevant literature. In Section 3, we characterize the cold-start problem in a simple model. In Section 4, we introduce the empirical model. In Section 5, we present the data. Section 6 and Section 7 contain the estimation procedure and results for the demand and supply side of the model respectively. In Section 8, we describe the model fit. In in Section 9, we describe the counterfactuals and present our counterfactual results. Section 10 concludes.

2 Literature

There is an extensive theoretical literature on social learning. Bergemann and Välimäki (1996, 1997, 2000) embed strategic experimentation into a market setting with endogenous prices. Che and Hörner (2018) and Kremer et al. (2014) employ a mechanism-design perspective and find that a recommender system which occasionally recommends "ex-ante unappealing" products can be "socially valuable because some of them are ultimately worthy of consumption" (p. 872, Che and Hörner, 2018). Hagiu and Wright (2020) analyze the platform-optimal level of experimentation which differs from the seller optimal level of learning if sellers have market power and the platform takes into account buyer surplus. Vellodi (2022) incorporates entry and exit of sellers. In his paper, consumers benefit if product ratings are censored because it alleviates the cold-start problem. We contribute to this literature by providing novel, empirical results about the scale and scope of the cold-start problem. Apart from our paper, to our knowledge, only Pallais (2014) assesses the cold-start problem empirically. In contrast to our paper, Pallais uses an experimental analysis in a labor market context.

Multiple studies estimate the impact of reviews on sales (Anderson and Magruder, 2012; Chevalier and Mayzlin, 2006), revenues (Luca, 2016), and exit rates (Cabral and Hortacsu, 2010). Additionally, several studies attribute substantial consumer surplus gains to online rating systems (Fang, 2022; Lewis and Zervas, 2016; Reimers and Waldfogel, 2021; Wu, Che, Chan and Lu, 2015). Bao, Fang and Osborne (2024) analyzes how quality disclosure via reviews affects entry and exit dynamics. We contribute to this literature by confirming the significance of ratings on Airbnb, demonstrating that unreviewed Airbnb listings are booked less frequently, are cheaper, and are more likely to be discontinued than reviewed ones.

A large body of empirical literature on Airbnb reviews examines the incentives for guests to write reviews and the sources of review bias (Fradkin and Holtz, 2022; Fradkin, Grewal and Holtz, 2021; Proserpio, Xu and Zervas, 2018; Zervas, Proserpio and Byers, 2021). Fradkin and Holtz (2022) highlight that, since consumers bear the cost of writing a review but do not receive all the benefits, online reviews are likely under-provided. Airbnb has also been the subject of extensive structural modeling. Most of these studies focus on the housing market (Almagro and Domınguez-Iino, 2024; Calder-Wang, 2021) or hospitality market (Farronato and Fradkin, 2022). Huang (2022) examines pricing behavior on Airbnb and finds that pricing frictions lead to a substantial welfare loss. Rossi (2024) explores the role of reviews in motivating sellers to exert effort, and examines how competition mediates this relationship.

We also contribute to the literature on estimating dynamic oligopoly models. Aguirregabiria, Collard-Wexler and Ryan (2021) provide an excellent recent overview. Most existing studies focus on oligopolistic games involving entry, exit, and sometimes innovation (Barwick and Pathak, 2015; Collard-Wexler, 2013; Igami and Uetake, 2020; Kellogg, 2014; Takahashi, 2015; Wollmann, 2018). Another strand in this literature incorporates dynamic considerations into pricing decisions. For instance, in studies by Huang (2022), Williams (2022) and Hortaçsu, Natan, Parsley, Schwieg and Williams (2023), dynamic pricing arises from intertemporal price discrimination for products with a natural expiration date. In other contexts, dynamic pricing reflects the notion that price not only influences current demand but also acts as an investment. This phenomenon can be driven by network effects (Dubé, Hitsch and Chintagunta, 2010), learning-by-doing (Besanko, Doraszelski and Kryukov, 2014, 2019; Besanko, Doraszelski, Kryukov and Satterthwaite, 2010; Sweeting, Jia, Hui and Yao, 2022), switching costs (Chen, 2016), entry deterrence (Sweeting, Roberts and Gedge, 2020) or learning about product quality (Ching, 2010).

Similar to Ching (2010), we consider a setting where firm price moderates the speed at which consumers learn about its product. However, Ching (2010) does not explicitly consider the inefficiency due to social learning and analyzes the market for prescription drugs. Unlike the above papers, we solve our model using oblivious equilibrium concept developed in Weintraub et al. (2008) and Weintraub, Benkard and Van Roy (2010). This is particularly suitable in our setting of many small firms. The oblivous equilibrium has been applied in other studies, such as (Bao et al., 2024; Brancaccio, Kalouptsidi and Papageorgiou, 2020; Chen and Xu, 2023; Frechette, Lizzeri and Salz, 2019).

3 The Cold-Start Problem

In this section, we analyze the cold-start problem in a simple model to motivate our counterfactual analysis and provide intuition for our results.

Suppose there are an entrant and an incumbent firm supplying products E and I respectively at zero marginal cost. Let the success rate or "quality" of product $j \in \{E, I\}$ be denoted by $\omega_j \in [0, 1]$. The quality ω_I of I is publicly known and reflects the probability with which consumption of I is a "success" and yields utility of one. If it is a "failure", its utility is zero. The quality ω_E of E is unknown to both consumers and firms, and distributed according to a beta distribution with parameters $a, b > 0$.

The model has two periods, i.e., $t \in \{1, 2\}$. In each period, a risk-neutral Bayesian consumer arrives to purchase E or I. Crucially, the consumer in $t = 1$ (consumer 1) is distinct from the consumer in $t = 2$ (consumer 2). The prior belief of the consumer 1 about E's quality is characterized in Equation (1).

$$\omega_{E1} \equiv \mathbf{E}_1[\omega_E] = \frac{a}{a+b}. \tag{1}$$

If consumer 1 chooses E, she truthfully reports her experience (success/failure) with probability $v_r \in (0, 1]$ to consumer 2. Note that consumer 1's experience follows a Bernoulli distribution with the success probability equal to ω_E. Conditional the consumer 1 writing a review, the posterior belief ω_{E2} of consumer 2 about the quality of E is $\frac{a+1}{a+b+1}$ in case of success and $\frac{a}{a+b+1}$ if consumer 1 experienced a failure.

The expected indirect utilities of product $j \in \{E, I\}$ are given in Equation (2). The unit price of product j is denoted by p_j and ϵ_{jt} represents the idiosyncratic taste shock.

$$u_{jt} = \omega_{jt} - p_{jt} + \epsilon_{jt} = \nu_{jt} + \epsilon_{jt}, \text{ where } \epsilon_{jt} \overset{iid}{\sim} Gumbel(0, \pi^2/6), \tag{2}$$

Under the distributional assumption on the taste shocks, consumer t chooses E with probability $q_{Et} = \exp(\nu_{Et}) / (\exp(\nu_{It}) + \exp(\nu_{Et}))$. The cold-start problem arises because v_{1E} does not account for the fact that the review that may be generated if consumer 1 purchases product E in expectation benefits consumer 2. All else equal, the probability that product E is purchased is inefficiently low as a result.

Firms, on the other hand, exist for both periods and take the effect of their pricing decision in $t = 1$ on their profit in $t = 2$ into account. In $t = 1$, firm j solves the profit maximization problem in Equation (3). Notice that the firm's profit in $t = 2$ depends on p_{E1} and p_{I1} as the prices determine the likelihood that E is bought (and reviewed) in $t = 1$.

$$\max_{p_{j1}} \left(q_{j1} p_{j1} + \delta \mathbb{E}_2 \left[q_{j2} p_{j2} | p_{E1}, p_{I1} \right] \right) \tag{3}$$

Equation (3) characterizes the prices in the subgame-perfect Nash equilibrium. π^i_{j2}, $i \in \{0, g, b\}$, denotes firm j's profit in $t = 2$ if its product receives no review (0), a good review (g) or a bad review (b) respectively. Furthermore, $\phi_0 = -1, \phi_g = \omega^e_{E1}$ and $\phi_b = 1 - \omega^e_{E1}$.

Lemma 1. *Suppose that firms compete in Nash-Bertrand fashion. In equilibrium, the entrant firm and the incumbent firm set prices p^*_{E1} and p^*_{I1} respectively, where*

$$p^*_{E1} = \frac{1}{1 - q_{E1}(p^*_{E1}, p^*_{I1})} - v_r \delta \sum_{i \in \{0,g,b\}} \phi_i \pi^i_{E2}$$

$$\text{and} \quad p^*_{I1} = \frac{1}{q_{E1}(p^*_{E1}, p^*_{I1})} + v_r \delta \sum_{i \in \{0,g,b\}} \phi_i \pi^i_{I2}.$$

p^*_{E1} strictly decreases while p^*_{I1} strictly increases in v_r and δ.

The proof of the lemma can be found in Appendix A.1. Notice that both the entrant's and the incumbent's profits are convex in ω_{E1} and, in expectation, information revelation benefits both the incumbent and the entrant. Therefore, the entrant lowers its price in $t = 1$ and raises the likelihood of a review in $t = 2$, countering the cold-start problem and speeding up the revelation of its quality. To the same effect, the incumbent raises its price and further increases the likelihood that E is reviewed. Lemma 1 implies that the price difference between entrant and incumbent is larger in absolute terms if social learning plays a larger role ($v_r \delta \uparrow$).

According to Lemma 1, both firms have an incentive to alleviate the cold-start problem, though if consumer choices are socially efficient remains unclear. Note that even absent any considerations of their future profit, firms will set different prices than is socially optimal in the first period; whoever has a higher expected quality will set a higher price, while, as there is no difference in marginal costs, it is socially optimal that prices are equal across products. In the Nash equilibrium, social learning increases the absolute difference between entrant and incumbent prices when E's quality is expected to be worse than I's, while the opposite is true when the entrant is expected to be better. Proposition 1 characterizes the price difference in the first period which fully alleviates the cold-start problem and shows how it compares to the first-period price difference in the Nash equilibrium. Denote the difference in expected quality as $\tilde{\omega}_1 = \omega_{E1} - \omega_I$, the equilibrium price difference as $\tilde{p}^*_1 = p^*_{E1} - p^*_{I1}$ and the socially optimal price difference as $\tilde{p}^s_1 = p^s_{E1} - p^s_{I1}$.[3]

[3] The socially optimal price difference in the first period takes into account the Nash equilibrium played in the

Proposition 1. *Suppose a social planner sets prices in the first period.*

(i) The socially optimal price difference in $t = 1$ is negative irrespective of $\tilde{\omega}_1$, i.e.,

$$\tilde{p}_1^s < 0 \qquad \forall \qquad \tilde{\omega}_1.$$

(ii) If $\sum_i \delta v_r \phi_i (\pi_{I2}^i - \pi_{E2}^i) < 1$, there exists $\widehat{\tilde{\omega}_1} < 0$ such that $\tilde{p}_1^s < \tilde{p}_1^* \quad \forall \tilde{\omega}_1 > \widehat{\tilde{\omega}_1}$.

The proof of the proposition is relegated to Appendix A.2. According to Proposition 1(i), the social planner chooses p_{E1} to be smaller than p_{I1} to incentivize consumer 1 to purchase E and review it, irrespective of the expected quality of E and the quality of I. Furthermore, Proposition 1(ii) specifies when the socially optimal price difference in the first period is lower than in a Nash equilibrium. To be able to say something meaningful, we need to impose regularity condition $\sum_i \delta v_r \phi_i (\pi_{I2}^i - \pi_{E2}^i) < 1$. This condition makes sure that the information gain from an additional review is not excessively valuable to the incumbent compared to the entrant. In other words, the incumbent profits must not be more convex than the entrant profits by an order of magnitude. Under this condition, we can show that the entrant's price relative to the incumbent's price is higher in the equilibrium than is socially optimal, as long as consumers believe that the entrant's product is not much worse than the incumbent's product. Hence, the entrant charges a price that is too high and the market suffers the cold-start problem. In contrast to the consumers, the entrant and the incumbent are forward-looking. However, they account only for their own future expected profit but not for the future expected consumer surplus in their pricing decision.

Proposition 1 illuminates the underlying causes of the cold-start problem. In addition, it provides the rationale behind our counterfactual analysis, in which we examine the welfare implications of a tax-subsidy scheme designed to change the price disparity between relatively new and established products. However, it is important to note that the simple model is lacking important features which may affect severity or even the existence of the cold-start problem. Apart from featuring just two products, the simple model does not capture firms' entry and exit decisions. Second, there is no outside option available to consumers. Incorporating these elements in the simple model would prevent us from deriving theoretical results, which is why we require a structural analysis with the more complex model outlined in the next section. Beforehand, we offer a brief informal discussion of the consequences of these features:

Endogenous entry and exit Changing relative prices and thereby shifting demand from one listing to another not only influences the speed of learning, but also impacts per-period profits. Socially optimal prices are different from individually optimal prices; consequently, if only individual firms were forced to adopt the socially optimal price, their profit would decrease. As all prices change, the overall effect on individual seller profits become ambiguous. Profit changes trigger changes in the incentives for firms to enter and exit. However, each firm is not necessarily equally valuable to consumers, hence optimal prices will be affected by considerations

second period.

of market participation. In particular, the social planner may exercise caution when imposing profit reductions on firms that are deemed valuable, which are often incumbents.

Outside option When consumers have the option to choose an outside option, both the price difference and the *price level* will matter for social welfare. In a logit model, firms possess market power arising from heterogenous preferences; thus, reducing the price level generally enhances market efficiency in any given period. However, a lower price level can adversely affect profits and thereby entry incentives while increasing exit, which harms consumers and negatively affects welfare, if entry and exit are inefficient. Whether the social planner prefers a higher or lower price level compared to the Nash equilibrium will depend on the relative importance of these channels. Note, however, that both these channels are orthogonal to diverting demand from incumbent to entrant and thus unrelated to social learning and the cold-start problem.

4 Model

In this section, we generalize the model from the previous section and apply it to the context of Airbnb. We assume that the true quality of each listing is unknown, but reviews serve as publicly observable signals of quality. As in the simple model, information is symmetric. That is, Airbnb guests and hosts have the same information and, therefore, have the same beliefs about listing quality at all times. The time horizon is infinite, and hosts maximize the discounted sum of future profits through their pricing, entry, and exit decisions.

Indirect utility and demand. – There exists a set $\mathcal{I}$ of guests, indexed by i, and a set $\mathcal{J}$ of listings, indexed by j. Let N_{jt} denote the number of reviews and K_{jt} the number of good reviews of listing j accumulated up to period t. Guests have unit demand. Let u denote a guest's indirect utility of renting listing $j \in \mathcal{J}_t$ in time period $t \in \{1, .., +\infty\}$.

$$u(p_{jt}, x_{jt}) = \gamma \frac{a + K_{jt}}{a + b + N_{jt}} + \beta_{l(j)} + (1 + f)\alpha p_{jt} + \xi_{jt} + \epsilon_{jt} = \nu(p_{jt}, x_{jt}) + \xi_{jt} + \epsilon_{jt} \tag{4}$$

γ reflects the utility value of having a successful stay with certainty, $\beta_{l(j)}$ is the intercept coefficient, which differs by listing type $l(j)$, and α is the rental price coefficient. p_{jt} is the rental price, f is the fee Airbnb levies on consumers, the structural error ξ_{jt} captures the unobserved (to the econometrician) listing characteristics, and ϵ_{jt} is an idiosyncratic taste shock. As in Section 3, a and b govern the prior distribution over listings' success rates, i.e., $\omega_j \overset{\text{iid}}{\sim} Beta(a, b)$. We allow for four different listing types, $l \in \{1, 2, 3, 4\}$, to account for observed differences in listing characteristics, such as a listing's location and amenities. A listing with more desirable characteristics is assigned a higher type. The details about how we construct listing types are found in Section 5. Notice that a listing's type and its rating are independent of each other. For example, there may be a type 1 listing with a bad rating and a type 4 listing with a good rating. The state x_{jt} of listing j in t is $(K_{jt}, N_{jt}, l(j)) \in X$ where $X = \{(K, N, l) \in \mathbf{N}_0^2 \times \{1, 2, 3, 4\} : K \leq N, N \leq \bar{N}\}$. $\bar{N} \in \mathbf{N}_+$ is the maximal number of reviews our model allows for.

As in Section 3, the taste shocks are independently and identically drawn from a normalized

Gumbel distribution. We also normalize the mean utility of taking the outside option, which can be interpreted as booking a hotel rather than an Airbnb, to zero. We abstract from guests arriving sequentially in the run-up to t to make bookings, as this would immensely complicate characterizing the demand system.[4] Rather, in each t, a discrete number of guests arrive simultaneously in the market to book accommodation for the duration of the period. Guest arrival follows a Poisson process with mean μ. In each t, an average number of μ guests arrive in the market to book one of the listings or take the outside option. Listings are capacity-constrained; at most, one consumer can rent listing j in t.[5] Equation (5) characterizes the demand q for j in t. It equals the probability that at least one of the guests arriving in t wants to book j. If more than one guest wants to book j in t, we assume that one guest successfully books the listing while all remaining ones are forced to take the outside option.

$$q(p_{jt}, x_{jt}, P_t, s_t) = 1 - \exp\left(-\mu \frac{\exp(v(p_{jt}, x_{jt}))}{1 + \exp(v(p_{jt}, x_{jt})) + \sum_x^X (s_t(x) - \mathbb{1}_{x=x_{jt}}) \exp(v(P_t(x), x))}\right) \tag{5}$$

$s_t(x)$ denotes the number of listings in each state x in t. Equation (5) assumes that hosts other than j set the rental price of their listings according to the pricing policy function $P_t(x)$. Listings that share the same state have the same price. This will be true in the symmetric equilibrium. Notice also that $P_t(x)$ does not depend on p_{jt} or the prices of $j' \neq j$. As discussed below, this follows from the oblivious equilibrium concept we use. Hence, we can simplify the notation and drop the j subscript in what follows.

State transitions. – If a listing is booked for t, with probability $v_r \in (0, 1)$, the guest accurately reports its experience (success, failure) in a review.[6] The probability ρ_0 that a listing's state does not change, either because it is not booked or the guest fails to leave a review, is $1 - v_r q$. The listing receives a good review with probability ρ_g or a bad review with probability ρ_b, depending on the listing's quality prior.

If the review is good, both N and K increase by one in $t + 1$. If the review is bad, N increases by one, but K does not. The possible transitions are illustrated in figure Figure 1.

Equation (6) summarizes the transition probabilities.[7]

$$\begin{aligned} \rho_0(p_t, x_t, P_t, s_t) &= 1 - v_r q(p_t, x_t, P_t, s_t) \\ \text{and} \quad \rho_g(p_t, x_t, P_t, s_t,) &= v_r q(p_t, x_t, P_t, s_t) \frac{a + K_t}{a + b + N_t} \\ \text{and} \quad \rho_b(p_t, x_t, P_t, s_t) &= v_r q(p_t, x_t, P_t, s_t) \left(1 - \frac{a + K_t}{a + b + N_t}\right) \end{aligned} \tag{6}$$

[4] If consumers arrive sequentially, expected demand does not have a simple, closed-form solution because the current set of available listings depends on past booking decisions. While it is possible to integrate the different booking sequences numerically, it is computationally expensive.

[5] We show in Online **??** that the number of consumers who want to book listing j is Poisson distributed.

[6] Note that we assume that reviews are truthful and the probability of leaving a review is exogenous. The incentive of consumers to leave a product review may be manifold and is the subject of analysis in many other scholarly contributions, such as Fradkin et al. (2021). Modeling this is beyond the scope of this paper.

[7] A full formulation of the transition matrix $\mathbf{T}(\mathbf{P}_t, \mathbf{s}_t)$, where $\mathbf{P}_t = P_t(\mathbf{X})$ and $\mathbf{s}_t = s_t(\mathbf{X})$ can be found in Online Appendix B.1.

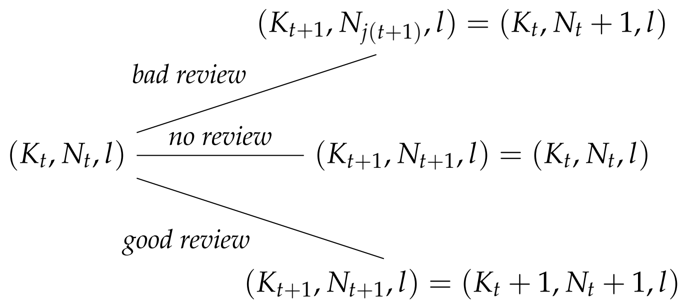

Figure 1: State transitions.

The host's problem. – There is a finite number of J potential hosts, but not all of them are active. Each active host operates a single listing.[8] An active host makes per-period revenue $\pi(p_t, x_t, P_t, s_t) = (\text{period length in days}) q(p_t, x_t, P_t, s_t) p_t$. We do not model a marginal cost of renting out a listing. Instead, a host incurs operating cost ϕ_{lt}, which is a fixed cost. ϕ_{lt} reflects that a host cannot use their apartment themselves for the time it is listed on Airbnb. This holds irrespective of whether the listing is booked or not. There is also a fixed cost κ_{lt} from entering the market and becoming active as the host must convert their apartment to an Airbnb rental. The operating and entry costs are random variables that vary over time. ϕ_{lt} and κ_{lt} follow exponential distributions with mean $\bar{\phi}_l$ and $\bar{\kappa}_l$ respectively. That is, $\kappa_{lt} \overset{\text{iid}}{\sim} Exponential(\bar{\kappa}_l)$ and $\phi_{lt} \overset{\text{iid}}{\sim} Exponential(\bar{\phi}_l)$. As the costs may differ for listings with different characteristics, we allow their means to vary across listing types.

Timing. – In each t the timing of events is a follows.

1. Active hosts set their prices and receive the per-period revenue.
2. Inactive hosts observe κ_{jt} and decide whether to enter at the beginning of $t+1$ or not. They pay κ_{jt} if they enter.
3. Active hosts observe ϕ_{jt} (the cost of operating the listing in $t+1$) and decide whether to exit at the end of t or not. They pay ϕ_{jt} if they do not exit.
4. Review outcomes are determined, and the industry takes on a new state $\mathbf{s}_{t+1}$.

Value function. – Having entered, in each t, the active host maximizes the expected discounted profit flow through her exit and pricing decisions.[9] We denote the discounted profit flow from the optimal pricing and exit behavior conditional on staying in the market in the current period by $V(x_t, P_t, s_t)$, the value function. Notice that V depends on the pricing policy function $P_t(x)$ and the number of listings $s_t(x)$ in each state. In equilibrium, all hosts behave according to a common, stationary strategy. This allows us to drop P_t and s_t as arguments of

[8] 90% of hosts in our dataset operate a single listing. The average number of listings per host is 1.10.
[9] Notice that we do not allow hosts to invest in quality, as listing quality is assumed to be exogenous. In principle, hosts could improve their services to guests to get better reviews. We find that ratings are highly correlated across categories, such as communication, accuracy, cleanliness, check-in, and location. Location-related ratings are likely exogenous.

V. We write V recursively in Equation (7). Note that the host shares a common prior with consumers, which precludes any form of price signaling. In practice, this means that the host is as good as any consumer when judging the quality of its listing.

$$V(x_t) = \max_{p_t} \left\{ \pi(p_t, x_t) + \mathbb{E}_\phi \left[\max\left(0, \delta \mathbb{E}_{x_{t+1}} \left[V(x_{t+1}) | p_t, x_t \right] - \phi_{lt} \right) \right] \right\} \tag{7}$$

$\delta \in (0,1)$ denotes the discount factor. Equation (7) incorporates the host's decision to exit the market when the current operating cost exceeds the expected value of remaining in the market. A host who exits the market cannot re-enter and is replaced with an inactive host.

Exit & entry rates. – Recall that a host exits the market if ϕ_{lt} exceeds $\mathbb{E}_{x_{t+1}} \left[V(x_{t+1}) | p_t, x_t \right]$ and that ϕ_{lt} is exponentially distributed. Then, the probability that an active host of a listing in state x_t exits is given by $\chi(p_t, x_t)$.

$$\chi(p_t, x_t) = \exp(-\delta \mathbb{E}_{x_{t+1}} [V(x_{t+1}) | p_t, x_t] \bar{\phi}_l^{-1}) \tag{8}$$

If a host has not entered the market, it is inactive. A listing is sure to have no reviews upon entry. An inactive host, therefore, chooses to enter if $\kappa_{lt} \leq V((0,0,l))$ and remains inactive otherwise. The entry probability of an inactive host of a type l listing is λ_l.

$$\lambda_l = 1 - \exp(-\delta V((0,0,l)) \bar{\kappa}_l^{-1}) \tag{9}$$

Using the fact that the operating cost is exponentially distributed, we specify hosts' exit behavior.[10] We rewrite Equation (7) as follows.

$$V(x_t) = \max_{p_t} \left\{ \pi(p_t, x_t) + \delta \mathbb{E}_{x_{t+1}} [V(x_{t+1}) | p_t, x_t] - (1 - \chi(p_t, x_t)) \bar{\phi}_l \right\}, \tag{10}$$

The *expanded* transition matrix $\mathbf{F}(\mathbf{P}_t, \mathbf{s}_t, \chi, \lambda)$ extends $\mathbf{T}$ to cover transitions from and to activity.[11] $\mathbf{F}$ allows us to compute the stationary listing distribution $\mathbf{s}$ across states implied by $\mathbf{F}^T \mathbf{s} = \mathbf{s}$.

Equilibrium. – Computing a Markov perfect equilibrium, as Ericson and Pakes (1995) do in their seminal paper on industry dynamics, is computationally infeasible in our context. In the Markov perfect equilibrium, the dimension of the state space increases exponentially in the number of hosts, and so does the cost of computing the Markov perfect equilibrium. It is prohibitively large for thousands of hosts.

Instead, we compute the oblivious equilibrium introduced in Weintraub et al. (2008).[12] The oblivious equilibrium can be interpreted as an approximation of the Markov perfect equilibrium of a game with a large number of players whose strategic response to a change in a single player's behavior is negligible. In the context of Airbnb, a host is unlikely to adjust its rental price in response to a price change by one of its thousands of competitors. Each host considers

[10] We show in Online Appendix B.4 that the expected operating cost is $(1 - \chi(x))\bar{\phi}_l - \mathbb{E}_{x_{t+1}}[V(x_{t+1}) | p_t, x_t] \chi(x)$.

[11] We provide a full formulation of $\mathbf{F}$ in Online Appendix B.1.

[12] There exist other equilibrium concepts for large dynamic games (Doraszelski and Judd, 2012, 2019; Fershtman and Pakes, 2012). For these, unlike for the oblivious equilibrium, the size of the state space depends on the number of players.

only its own state and the stationary equilibrium listing distribution $\mathbf{s}$ when choosing its pricing strategy $P(x)$. In equilibrium, $\mathbf{s}^*$ is consistent with the hosts' pricing, exit, and entry behavior.

Let $\mathbf{s}^*$ be the stationary listing distribution that arises from each host behaving according to $\mathbf{P}^*$. $\mathbf{P}^*$ is the oblivious equilibrium strategy if and only if no host has a strict incentive to deviate from P^* to any P' in strategy set Π if all other hosts behave according to $\mathbf{P}^*$ and $\mathbf{s}^*$ remains unaffected by the host's deviation. Equation (11) formalizes this notion. Intuitively, each host behaves like an atom whose deviation $\mathbf{s}^*$ does not respond to. It is important to realize that hosts have market power regardless, as listings are differentiated.

$$V(x|P^*, P^*, \mathbf{s}^*) \geq V(x|P', P', \mathbf{s}^*) \;\forall\; P' \in \Pi. \tag{11}$$

5 Data

Airbnb is the market-leading peer-to-peer platform for short-term accommodation. It enables hosts to rent their apartment or a room to guests, usually tourists. Since its founding in 2008, Airbnb has grown to feature more than four million Airbnb hosts worldwide, housing about 33 million guests annually on average.[13] Airbnb's annual revenue of roughly 6 billion US dollars in 2021 rivals the revenue of large hotel chains. With thousands of active Airbnb listings, New York City is by far Airbnb's largest market among US cities. Due to many *a priori* unknown listings, the crowd-based rating system of Airbnb is one of its key services to potential guests. Guests are encouraged to leave a review of the listing on a one-to-five stars scale within two weeks after concluding their stay, which is then published for other potential guests to see. The average star rating and the cumulative number of reviews are prominently displayed on Airbnb's website.

We use data collected by AirDNA, a data analytics company, on all Airbnb listings in Manhattan, New York, between January 2016 and December 2019.[14] Figure 2 (left) illustrates the spatial distribution of listings across Manhattan. During this time, the number of listings is relatively stable, supporting the idea that the market is in a stationary equilibrium (see Figure 2, right). We chose the timeframe to exclude the period affected by the COVID-19 pandemic. We have rental prices in USD for each listing and day during the observation period and information on whether a listing was available, reserved, or blocked for private use by the host on a given day. In addition, we observe the number of reviews and the rating on a one-to-five stars scale for each listing in a roughly bi-monthly frequency.[15] Lastly, the dataset includes various listing attributes and amenities, which we use to define the relevant market. We focus on listings that offer the entire apartment for rent (as opposed to a single room), include at least one picture of the place, permit at most two guests and no pets, and feature one bedroom and one bathroom. We drop observations if they occur before the first booking of a listing because

[13] See `https://news.airbnb.com/about-us/`.

[14] See `https://www.airdna.co/`.

[15] The frequency of the ratings and reviews depends on the frequency of the AirDNA scraping. All other variables are available at the daily level. There are less than two weeks between 40% of review observations. Only 1% observations lie more than three months apart.

we suspect that the earliest observation often predates the actual market entry. This way, we exclude listings that are never booked during their lifespan. We also drop listing-days with observed prices below the first and above the 99th percentile, corresponding to \$65 and \$518, respectively. We account for the cleaning fee hosts typically charge per stay by adding it to the daily price, divided by the average reservation length of 5.5 days. If we do not observe the cleaning fee for a listing, we assume it is equal to the average cleaning fee in the data sample.[16]

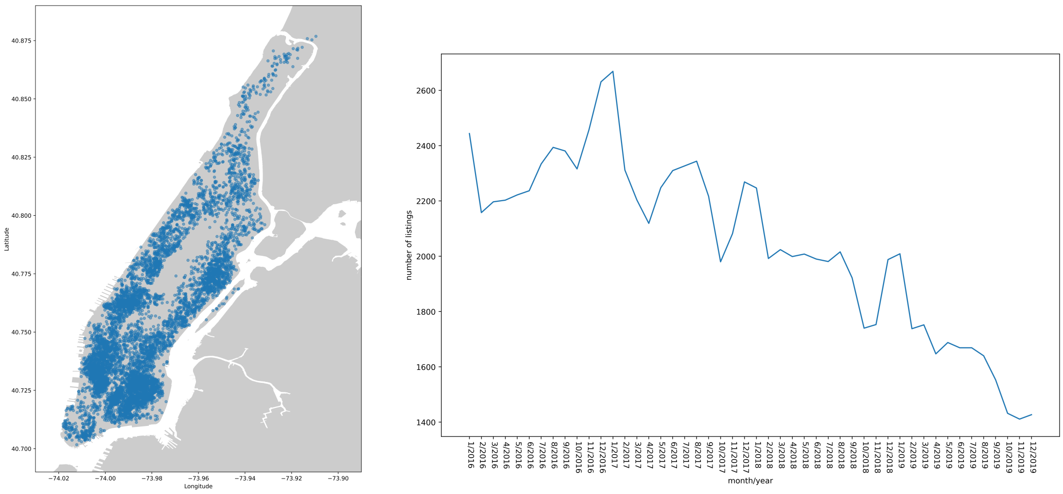

Figure 2: Number of listings over time and their location.

To estimate the model, we aggregate the data to four-week periods, or what we refer to as "months" in the following. This is because inversion of the demand system in the context of demand estimation precludes occupancy rates of 0% and 100%. After the previously described cleaning, we have 62,937 listing-month observations.

	mean	std	min	25%	50%	75%	max
Rental rate	\$193.02	\$60.13	\$70.33	\$150.75	\$184.78	\$270.95	\$562.43
Occupancy rate	60.64%	33.57%	0.00%	33.33%	69.23%	100.00%	100.00%
Number of reviews	10.30	8.19	0.00	2.00	9.00	20.00	20.00
Rating	4.51	0.72	1.00	4.40	4.67	5.00	5.00
Monthly exit rate	3.21%	0.80%	1.57%	2.68%	3.21%	4.08%	5.55%
Monthly entry rate	4.40%	1.85%	0.39%	3.17%	4.36%	6.26%	9.83%
Lifespan (in months)	17.67	16.38	1.00	3.00	12.00	39.00	52.00

Table 1: Data summary.

We construct the number of "good" reviews K as the number of five-star reviews required to achieve the observed rating if the remaining $N - K$ reviews were "bad" one-star ones. For example, for a listing with ten reviews and a rating of 3.8 stars, the implied number of good reviews is 7. We deal with missing information on the number of (good) reviews by filling in

[16] This leaves us with 7,586 unique listings and 2,298,598 listing-days, excluding blocked ones.

the most recent observed value. To keep a reasonably sized state space, we censor the number of total reviews at 20, i.e., $\bar{N} = 20$. Since the marginal effect of N on the posterior belief becomes small as N grows large, the change in the posterior mean tends to be minor for N exceeding 20. For example, suppose the prior quality distribution has a mean of 4.15 stars and a variance of 0.27.[17] Figure 3 illustrates how the posterior mean experiences ever smaller changes as the listing receives either a sequence of good or bad reviews. Adding a good review to a yet unreviewed listing improves its expected quality by 4.42%, whereas the expected quality of a listing with 20 good reviews would only move by 0.19%.

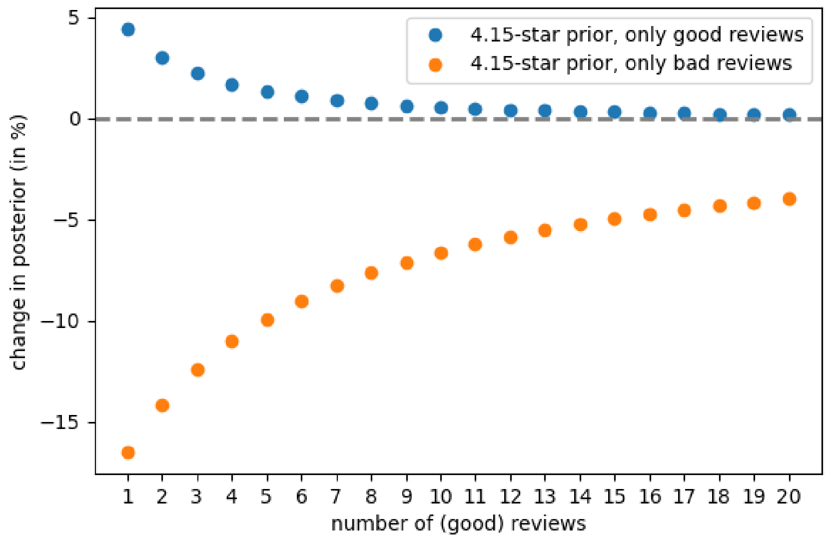

Figure 3: Example posterior changes.

Table 1 summarizes key variables in the data on the listing-month level. The average rental price is $193.02 in our sample. Rental prices exhibit little variation. The majority of listings are priced in the $50 range around the average price. On average, Airbnb listings are occupied 60.64% of the time and have been reviewed about nine times (after censoring the number of reviews). On average 3.21% of hosts leave the market each month and 4.40% have just entered.[18] The average lifespan within our sample is 1.3 years, but most listings exist for less than a year.[19]

At 4.5 stars, the average rating is high. In comparison, at 3.8 stars, the average rating on TripAdvisor is substantially lower (Zervas et al., 2021). It is a priori unclear if this is mainly because highly rated listings remain in the market longer or because the listings' quality is generally high. Figure 4 shows the rating distribution in our sample. The distribution is left-skewed because most listings are rated four stars or higher. Listings with a lower rating tend to have fewer reviews. Almost all listings with a 1-star, three-star, or 3.66-star rating have less than four reviews.

Conditional on the rental price and the reviews, we want listing types to capture as much

[17] These values correspond to our estimates (see Section 6).

[18] We use the whole sample with dates before January 2016 and after December 2019 to determine entry and exit. Rates are relative to the active hosts in the market.

[19] Here, the lifespan are the number of months for which we observe a listing in our sample, hence it has a natural maximum of 4 years.

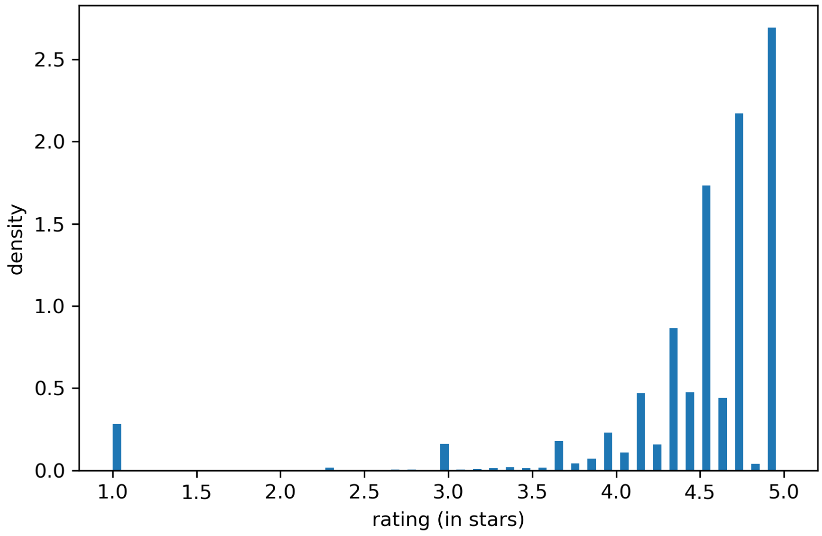

Figure 4: Rating distribution in the data.

heterogeneity in the utility that consumers derive from different listings as possible, both unobserved and observed. We choose to be parsimonious and distinguish only four types, thereby limiting the size of the state space and keeping the model tractable. To construct the four listing types, we regress the occupancy rate on the rental price, *K*-*N* fixed effects, month-year fixed effects, and listing fixed effects. We divide the estimated listing fixed effect coefficients into type 1, type 2, type 3, and type 4, depending on the quartile they fall into. Figure 5 shows how the type of a listing relates to the fixed effect we estimate for that listing and the listing's rental price.

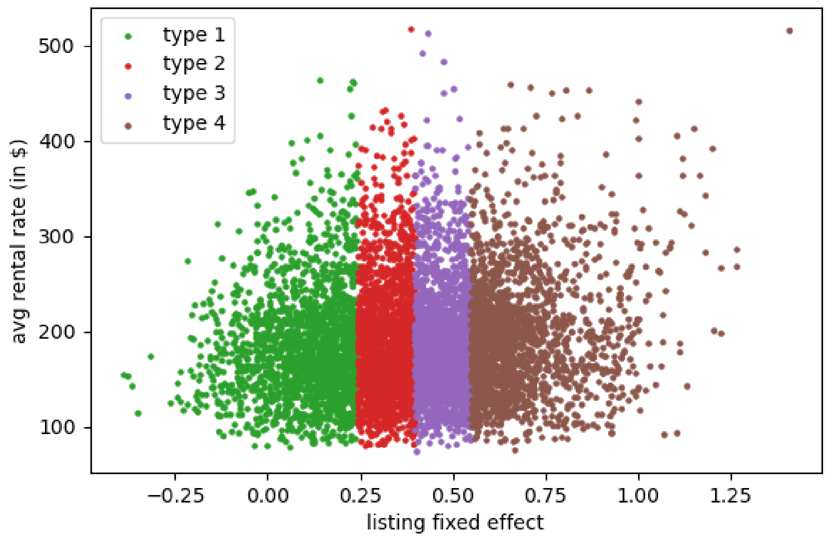

Figure 5: Listing type definition.

We report summary statistics for the different types in Table 2. On average, higher types feature a higher occupancy rate. While a type 4 listing is, on average, booked 81% of the time, the average occupancy rate of a type 1 listing is only 35%. The correlation between the type and the number of reviews is very weak (Pearson correlation coefficient: -0.013). The rating tends to increase in the type (0.09). As a results, also the rental price is positively correlated with the listing type (0.08). Higher-type listings also tend to be older. The correlation coefficient of a listing type and the lifespan is 0.11.

	Price	Occupancy	Reviews	Rating	Lifespan
type 1	\$186.30	34.60%	9.89	4.39 stars	14.6 months
type 2	\$192.83	53.47%	10.48	4.48 stars	20.2 months
type 3	\$189.45	68.58%	11.04	4.56 stars	21.2 months
type 4	\$204.40	80.92%	9.32	4.58 stars	14.7 months

Table 2: Summary: listing types.

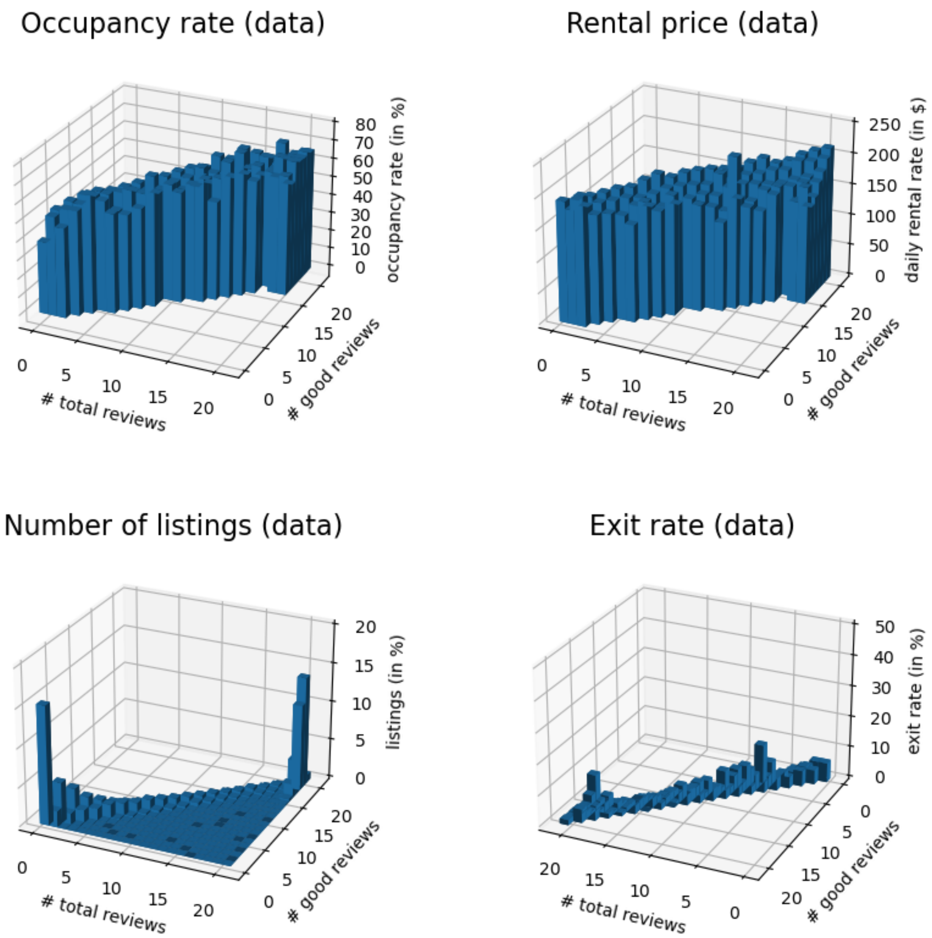

Figure 6: Occupancy rate (top left), rental rate (top right), number of listings (bottom left), and exit rate (bottom right) by number of (good) reviews in the data.

We plot the average occupancy rates, rental prices, number of listings, and exit rates depending on N and K in Figure 6. Occupancy rates tend to be larger for listings with many reviews, assuming these reviews are good (Figure 6, top left). Listings without reviews, for example, are around 40% less likely to be booked than listings with 20 good reviews. Similarly for rental prices, though the relationship is less stark. Listings with 20 good rather than zero review have a 12% higher price on average.[20]

Hosts of listings with many reviews and a good rating make more revenue and remain in the market, whereas they leave the market otherwise. Hence, our dataset contains very few or no observations for many states associated with small K, while most observations are concentrated on high-K states (see Figure 6, bottom left). The right-skewed listing distribution of firms over (good) reviews is driven by a selection effect reminiscent of Jovanovic (1982). High-quality listings survive, whereas low-quality ones fail and exit. The listing distribution features a "pitch fork" shape, as most listings either have few or no reviews or – due to right-tail compression – have the maximum amount. Notably, there are roughly 5.5 times as many reviewed listings as unreviewed ones. About 39% of reviewed listings have the maximal review count. We also observe the selection effect in hosts' exit behavior (see Figure 6, bottom right). Compared to a reviewed listing, an unreviewed listing is around three times as likely to exit in a given month.

To support the graphical intuitions of Figure 6, we regress various dependent variables on the rental price, the number of reviews, the number of good reviews, their interaction, listing types as well as year-month. In this way, our regression results account for differences in time-invariant listing characteristics and seasonality. They are reported in Table 3. Considering an average listing with ten reviews and a 4.5 star rating as baseline, a listing with one more review has a price that is higher by \$0.98 if the review is good and a price that is lower by \$10.55 if the review is bad. Comparing the same listings, the occupancy rate increases by about 1.33 percentage points and decreases by 1.08 percentage points after receiving a good and a bad review, respectively. Notice that the occupancy rate barely responds to an increase in the rental price by \$1 – it is a mere 0.1 percentage points lower in response. The exit rate is generally lower for a listing with an additional review, whether the review is good or bad seems to play a minor role. This could be the result of the strong selection effect in the dataset.

We do not estimate all the parameters of our model, but we calibrate some of them using data moments. Recall that a period is a four-week interval. We estimate the four-week review rate v_r, i.e., the number of additional reviews a listing has four weeks after being booked, as the change in the number of reviews over the number of bookings of listings with less than 20 reviews in our data sample. In this way, we determine that v_r is 0.992. Given that the average reservation length is 5.5 days, the probability that a guest leaves a review after the stay is around 20%. It is important to note that v_r does not depend on the number of (good) reviews, the rental price, or whether the guest experiences a success or a failure. Each booking is equally likely to result in a review. At any given time, we observe, on average, 1,210 active listings, of which 735 are booked. We set the mean number of guests arriving in the market to 10,000. Assuming that one in three consumers looking for accommodation considers Airbnbs, this implies that Airbnb's mean market share is 2.1%, consistent with the 2% median market share across 50 US cities in Farronato and Fradkin (2022). We set the number of potential listings J to 10,000. For

[20] Moreover, prices of entrants are relatively high, higher than the price of a listing with 20 reviews and a 4.6-star rating.

	Dependent variable:		
	rental rate	occ rate	exit rate
	(1)	(2)	(3)
rental rate		-0.001^{***}	-0.0001
		(0.000)	(0.000)
no of reviews	-10.340^{***}	-0.002^{***}	-0.003^{***}
	(0.105)	(0.004)	(0.001)
no of good reviews	11.795^{***}	0.035^{***}	-0.001
	(0.492)	(0.002)	(0.001)
no of reviews $\times$	-0.024	-0.001^{***}	0.0001^{***}
no of good reviews	(0.018)	(0.000)	(0.000)
Type FE	Yes	Yes	Yes
Observations	62,937	62,937	62,937

Note: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parenthesis clustered at type level.

Table 3: Rental rate, occupancy rate, exit rate and reviews in the data.

the discount factor δ, we choose 0.995, implying an annualized interest rate of 6.7%. Finally, Airbnb charges – both in the model and in reality – a commission fee of 14.2%.[21] The calibrated parameters are summarized in Table 4.

Parameter		Value
Discount factor	δ	0.995
Revenue fee	f	0.142
Arrival rate	μ	10,000
Review rate	v_r	0.992
Maximum number of reviews	$\bar{N}$	20
Maximum number of listings	J	10,000

Table 4: Calibrated parameters.

6 Demand Estimation

We estimate the parameters governing demand by the generalized method of moments. Some listings still have zero bookings in a given month (or, more rarely, are fully booked out), in which case we have to drop those observations for the demand estimation. This leaves us with 49,214 observations and 7171 unique listings. We invert demand and back out ξ_{jt} to compute the moment conditions. Equation (12) characterizes the regression equation we estimate.

$$\ln\left(\frac{-\ln(1-B_{jt})}{\mu_t}\right) - \ln\left(\frac{1+\sum_j^{\mathcal{J}_t}\ln(1-B_{jt})}{\mu_t}\right) = \nu(p_{jt}, x_{jt}) + \xi_{jt}. \tag{12}$$

We account for seasonality by allowing the arrival rate to vary by month. We compute μ_t as μ times the average percent deviation of the total number of Airbnb bookings from the mean in a given month. For example, in the first four weeks of the year (i.e., in January), the arrival rate is 22.8% lower than average. Following Ferrari and Cribari-Neto (2004) and Dickstein (2021), we do not estimate a and b directly, but estimate ψ and ι instead. They are defined in Equation (13). ψ determines the prior mean, whereas ι is closely related to the variance of the prior distribution.

$$\frac{1}{1+\exp(-\psi)} = \frac{a}{a+b} \quad \text{and} \quad \exp(\iota) = a+b. \tag{13}$$

This alternative formulation facilitates the estimation but also restricts the set of possible solutions. Naturally, the type coefficients and the coefficient of the rental price are identified by variation in the listing types and the rental price, respectively. The identification of γ, ψ, and ι comes from within- and between-listing variation in the number of (good) reviews. ι pins down to what extent an additional review moves the posterior mean away from the prior mean. If ι is small, the prior belief is precise, and guests make only marginal adjustments to their beliefs

[21] See `https://www.airbnb.ca/resources/hosting-homes/a/how-much-does-airbnb-charge-hosts-288`.

after observing the rating. ψ depends on variation in the rating to be identified. The posterior mean responds more strongly to reviews if the rating differs greatly from the prior mean. In particular, the posterior mean increases by more as the rating improves if the prior mean is low. In this way, the relationship between the rating and the posterior mean allows us to estimate the prior mean. γ captures the impact of the posterior mean on guests' booking decisions. If the reviews have little effect on the occupancy rate of the listings, regardless of their current rating and review count, γ must be low.

Column (1) in Table 5 shows the estimation results assuming independence of the structural error regarding the rental price and the number of (good) reviews. Clearly, this assumption may be violated in reality. First, the four listing types may not fully capture the time-invariant listing characteristics. In this case, the structural error is correlated with the occupancy rate and the number of reviews, which result from past bookings. Second, rental prices may be correlated with time-varying unobserved demand shocks. Third, hosts may exert effort to improve their listing's desirability and their incentive to do so may depend on market conditions. Using suitable instrumental variables, we account for the potential endogeneity of the rental price and the number of (good) reviews.

We instrument p with the average reservation length of a listing. The reservation length serves as a proxy for any cost of welcoming new guests (we have to abstract from this cost in the model). Properties occupied for a longer period by the same guests are less costly to maintain, which allows the host to charge a lower price. If the average reservation length is unrelated to a listing's characteristics, it satisfies the exclusion restriction. We instrument the number of (good) reviews with the de-meaned five-, six–, seven–, and eight-month lagged occupancy rates, their squares, and their interaction with the rating. Intuitively, an unexpectedly higher occupancy rate in the past has no bearing on the present occupancy rate other than through the reviews that resulted from it. The results of estimating the demand coefficients using instruments are reported in column (2) in Table 5.

As expected, the standard errors of coefficients that are identified with the instruments are relatively larger. The coefficient of the posterior mean narrowly fails to be significant on the 10% level. ψ and ι imply a prior mean and variance of 4.15 stars and 0.27, respectively. Figure 7 relates the estimated prior distribution to the rating distribution in the data. Recall that the rating is an imprecise signal of the true quality and features a larger variance. Also, the rating is subject to the selection effect; highly rated listings are observed in the data, whereas poorly rated listings are not.

The magnitude of the rental price coefficient is four times larger in column (2) compared to column (1), suggesting that the rental price is indeed endogenous in column (1). The estimated rental price coefficient corresponds to an average listing-month price elasticity of -1.04. To compare, Farronato and Fradkin (2022) report an average price elasticity of the demand for Airbnb and hotel accommodations of -4.27. In Huang (2022), the average price elasticity of demand for Airbnb accommodation in San Francisco is -2.51. While our estimated elasticity is somewhat lower in absolute terms compared to the literature, we believe it is reasonable. It is possible that our estimate of the price coefficient reflects search frictions; if guests are not aware of all avail-

		(1)	(2)
prior	ψ	-0.1810	1.3171**
		(0.1463)	(0.7421)
	ι	1.9633***	1.6212
		(0.1233)	(1.1995)
rental rate	α	-0.0020***	-0.0086***
		(0.0001)	(0.0015)
types	β_1	-10.6106***	-10.5354***
		(0.1109)	(1.7158)
	β_2	-10.1166***	-9.8218***
		(0.1110)	(1.7193)
	β_3	-9.6580***	-9.4401***
		(0.1116)	(1.7221)
	β_4	-9.2443***	-8.9099***
		(0.1114)	(1.7415)
expected quality	γ	1.6000**	2.8607
		(0.1358)	(1.8711)
Observations		49,214	25,824

Note: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parenthesis.

Table 5: Demand estimates.

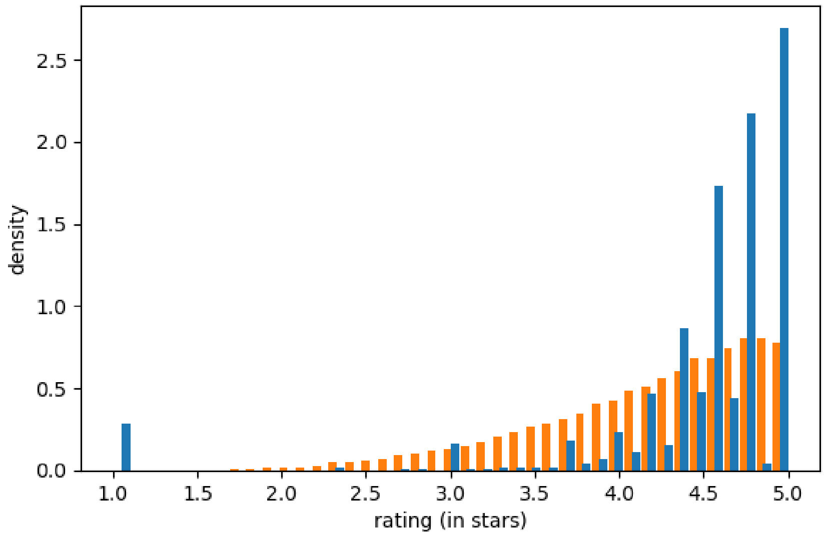

Figure 7: Estimated prior distribution and rating distribution in the data.

able listings, they would substitute to other options at a comparatively lower rate in response to a price increase, leading to relatively low price elasticity.

	mean	std	min	25%	50%	75%	max
Own-price elasticity	-1.0426	0.7294	-5.4299	-1.4287	-0.9022	-0.4941	-0.0005
Good review semi-elasticity	2.5640	2.7082	0.0041	0.6497	1.3687	3.6880	16.6112

Table 6: Summary of elasticities.

We also compute the semi-elasticity of demand if a listing receives a good review. On average, the occupancy rate increases by 2.56% in response to a good review. Figure 8 shows that the demand for listings with fewer reviews is more elastic. Guests who see an additional review will adjust their expectations of a listing's quality only slightly if the listing already has many reviews and the rating is precise. Conversely, observing a review has a much greater impact on guests' beliefs if the listing has few or no reviews and there was previously little information available to assess its quality. Accordingly, the average good review semi-elasticity is 5.80 for unreviewed listings but a mere 0.68 for listings with 19 reviews. Table 6 provides summary statistics for the two elasticities we estimate.

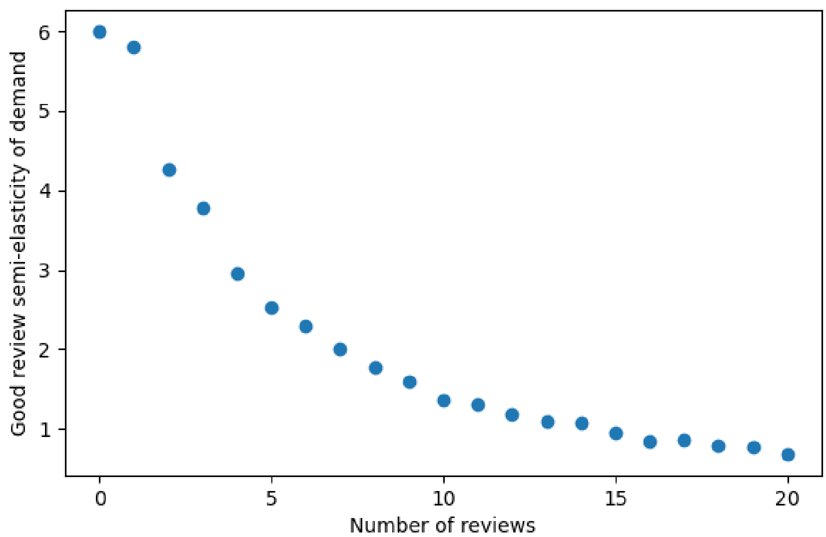

Figure 8: Mean good review semi-elasticity conditional on the number of reviews.

The prior distribution parameter estimates, together with the estimates of the expected quality and the rental price, imply that guests' willingness to pay is *ceteris paribus* $11.56 higher per day for a listing with one review and a 5-star rating compared to an unreviewed listing. By contrast, an additional good review of a listing with 19 reviews and a four-star rating is worth only $3.34 per day to guests.

Recall that Airbnb has a low single-digit market share. The type coefficients are estimated to be negative and large in magnitude, reflecting that most guests take the outside option (i.e., book a hotel) rather than book an Airbnb. Remember that a listing is attributed a higher type if it has a higher occupancy rate after accounting for the rental price and the number of (good) reviews. Hence, the type coefficients increase in the type.

7 Supply Estimation

We estimate the cost parameters of the model, denoted by c, conditional on the demand parameter estimates $\theta = \{\psi, \iota, \alpha, \beta_{,1}, \beta_2, \beta_3, \beta_4, \gamma\}$, by Maximum Likelihood. The log-likelihood of the number of listings across states is depicted in Equation (14). $s^*(x|c,\theta)$ denotes the number of listings in state x in the oblivious equilibrium of our model. y_{jt} is a variable indicating the state of listing j in month t. $s^*(x|c,\theta)$ is determined by the demand parameters (as they determine how frequently listings change states) and the cost parameters (because the listings' entry and exit rates depend on them). Conditional on the demand parameters, the operating costs, and the entry costs are identified by variation in the number of listings across states. We assume that listings are equally likely to have any of the four types. Hence, the average number of listings of a certain type outside the market is 2,500 minus the total number of listings of that type in the market. Denote the set of states associated with type l by X_l

$$\begin{aligned} L(c|y_{jt},\theta) = & \sum_t \left\{ \sum_j \left[\sum_{x \in X} \mathbb{1}(y_{jt} = x) \ln(s^*(x|c,\theta)) \right] \right. \\ & \left. + \sum_{l \in \{1,2,3,4\}} \left(\frac{J}{4} - \sum_{x \in X_l} \sum_j \mathbb{1}(y_{jt} = x) \right) \ln \left(\frac{J}{4} - \sum_{x \in X_l} s^*(x|c,\theta) \right) \right\} \end{aligned} \tag{14}$$

The maximum likelihood estimation requires us to solve the model and determine $\mathbf{s}^*$ repeatedly for different cost parameter candidates. We solve the model in four steps. We start by formulating the initial guess of the pricing policy function $P_0(x)$, the number of listings $s_0(x)$ per state, and the host's value function $V_0(x)$. Initially, half of the hosts are active and half of them are inactive. All hosts charge a price of $200 and their value function is the present discounted value of their revenue. The number of listings is identical across states.

Step 1 – Based on the guess, we determine a host's best response $P_1(x)$ if all remaining hosts adhere to $P_0(x)$. We solve for $P_1(x)$ using Newton's method. Specifically, we iterate over Equation (15), where k is the iteration step and $P_1^0 = P_0(x)$, until the change in the rental price, $P_1^{k+1} - P_1^k$ for any state x is less than $0.1.

$$P_1^{k+1}(x) = P_1^k(x) - \frac{v'(x)}{v''(x)} \tag{15}$$

Functions v' and v'' are the respective first- and second-order derivatives of the host's value function with respect to the rental price.[22]

Step 2 – Assuming that all hosts set their prices according to $P_1(x)$, we compute the occupancy rate $q_1(x)$ and transition matrix $T_1(x)$. We use them to compute the value function $V_1(x)$ from Equation (10).

Step 3 – We use $V_1(x)$ to update the exit rate in state x to $\chi_1(x)$ and the entry rate for each type l to λ_l. Together with the occupancy rates $\mathbf{q}_1$, χ_1 and $\boldsymbol{\lambda}_1$ allow us to compute the expanded transition matrix $\mathbf{F}_1$. We use $\mathbf{F}_1$ to solve for the new, stationary listing distribution $\mathbf{s}_1$, where $\mathbf{s}_1$

[22] The exact expressions can be found in the Appendix B.2.

is given by:

$$\mathbf{s}_1 = \mathbf{s}_1 \mathbf{F}_1 \tag{16}$$

Step 4 – If the absolute difference between $V_1(x)$ and $V_0(x)$ or P_1 and V_0 or $s_1(x)$ and $s_0(x)$ for any x exceeds 0.000001, we update the guess to $P_1(x), s_1(x)$, and $V_1(x)$ for all x and repeat steps 1 to 3. Otherwise, $(\mathbf{P}_1, \mathbf{s}_1, \mathbf{V}_1)$ constitutes the model solution $(\mathbf{P}^*, \mathbf{s}^*, \mathbf{V}^*)$.

Doraszelski and Satterthwaite (2010) establish the existence of a symmetric equilibrium in pure strategies for a closely related model. Note that the model may have multiple equilibria. (Doraszelski and Satterthwaite, 2010; Weintraub et al., 2008).

		(1)
mean entry cost	$\bar{\kappa}_1$	181,368***
		(3,946.58)
	$\bar{\kappa}_2$	264,819***
		(5,136.71)
	$\bar{\kappa}_3$	426,997***
		(8,345.07)
	$\bar{\kappa}_4$	796,930***
		(11,485.67)
mean operating cost	$\bar{\phi}_1$	2,323***
		(1.40)
	$\bar{\phi}_2$	3,587***
		(2.02)
	$\bar{\phi}_3$	4,345***
		(2.79)
	$\bar{\phi}_4$	5,572***
		(3.19)
Observations		62,937

Note: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parenthesis.

Table 7: Supply estimates.

The estimation results are shown in Table 7. All estimated parameters are highly significant. Our estimates imply that hosts on average incur a $3,065 cost upon entry.[23] This compares to the average entrant's lifetime profit of $5,940 in present discounted value terms. Listings of any two types are by construction equally likely observed in the market although higher type listings are more profitable. Hence, the entry costs increase in listing type. Our estimates suggest that the next higher type requires 30 to 75% higher entry costs. For instance, the entry costs of a type-3

[23] Note that the average incurred entry cost is by many orders of magnitude lower than the average entry cost draw as firms will only enter and incur the cost if the draw is low enough.

listing are on average $898 or 33% higher than a type-2 listing.

We estimate that on average hosts pay $2,584 in operating costs. To compare, the average revenue per month is $3,422. This implies an average profit margin of 24.5%. As with entry costs, higher-type listings incur higher operating costs. Intuitively, higher-type listings are more desirable to guests but also the host. The host's opportunity cost of renting out the apartment, rather than using it herself, is therefore higher. A type-4 listing, for example, costs on average $3,656 or 173% more to maintain than a type-1 listing. Nonetheless, hosts earn more profit from higher-type listings. The present discounted values of type-1, type-2, type-3, and type-4 listings are $5,213, $8,510, $11,481, and $13,235 respectively.

8 Model Fit

We use our estimates to simulate four years worth of data. We compare the simulated data to the actual data to assess the fit of our model. Table 8 shows key simulated data moments. As is to be expected, most variables exhibit less variation in the simulated data compared to the real data. The average rental price is roughly 5% lower than in the actual data. The average occupancy rate in the simulated data is almost identical to the one in the real data. Nonetheless, the simulated listings have on average fewer reviews (7) than those in the actual data (10). This is likely because hosts tend to exit at a relatively higher rate in the simulated data. The average monthly exit rates are 12.88% and 3.21% in the simulated and actual data respectively. The average rating in the simulated data is 4.63 stars, which is comparable to the 4.51-star average rating in the actual data.

	mean	std	min	25%	50%	75%	max
Rental rate	$182.58	$31.53	$119.31	$160.41	$182.07	$209.10	$248.22
Occupancy rate	60.01%	48.99%	0.00%	0.00%	0.00%	100.00%	100.00%
Number of reviews	7.32	7.60	0.00	1.00	4.00	14.00	20.00
Rating	4.63	0.86	1.0	4.71	5.00	5.00	5.00
(Monthly) exit rate	12.88%	0.33%	0.00%	0.00%	0.00%	0.00%	100.00%

Table 8: Simulated data summary.

Figure 9 compares the rating distribution in the simulated data ("rating (model)") to the rating distribution in the actual data ("rating (data)"). The frequency of one-star, three-star, $3\frac{2}{3}$-star ratings (i.e., ratings comprised of two good reviews and one bad review), and four-star listings (i.e., ratings comprised of three good reviews and one bad review) are remarkably similar in the simulated and the actual data. The model predicts more five-star ratings than we find in the actual data. In the actual data, listings at the high end of the distribution typically have ratings between four and five stars. We believe that the difference can be explained by the fact that the number of reviews is censored at 20. Hence, the model allows for little variation in the ratings close to 5 stars. Ratings in the simulated data are based on at most 20 reviews but possibly on many more in the actual data. Note that Figure 9 illustrates the selection effect that underlies the simulated and the actual data. Recall that the mean quality of *entrant* listings is estimated to be 4.15 stars. This is lower than the average rating of *all active* listings, both in the simulated and actual data. the 25th percentile is also fairly similar in the simulated and actual data. Less than 25% of listings have a quality of around 4.5 stars or less.

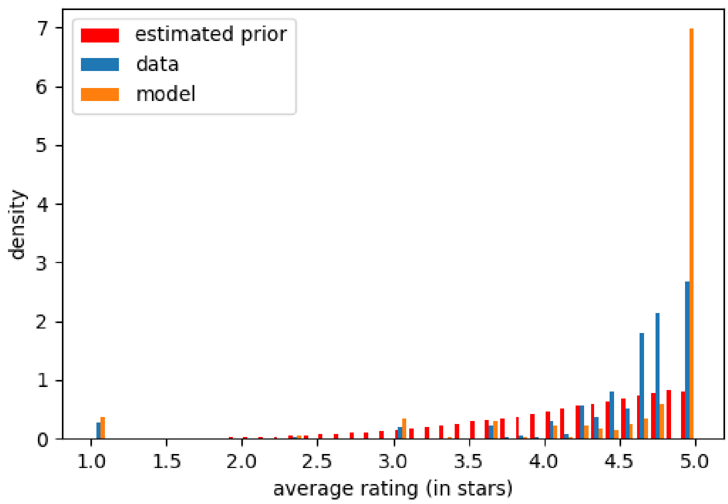

Figure 9: Estimated prior distribution, rating distribution in the data, and model rating distribution.

Recall that we estimate the cost parameters of the model by maximizing the likelihood of the equilibrium state distribution. According to our model, in equilibrium, there are on average 1,152 active listings, compared to 1,210 in the data. Figure 10 (bottom left) shows that the number of listings generated by the model for each number of (good) reviews closely matches the observed data. As in the actual data, around 20% of active listings are unreviewed. The model predicts that there are few to no poorly rated listings, especially for states associated with a relatively large number of reviews. Highly rated listings, on the other hand, tend to remain in the market and gather reviews. The model gives rise to a "pitched-fork" shaped distribution which is consistent with the data. As discussed, the model produces rental prices and occupancy rates that are comparable to the actual data (Figure 10, top right and top left, respectively).

It is critical for our counterfactual analysis that we get the behavior of rental prices, occupancy rates, and exit rates in the number of (good) reviews right. Due to the selection effect, we do not have data for many states. For the states we do have data for, the occupancy rate seems to respond more strongly to the accumulation of reviews compared to the model. Changes in rental prices, on the other hand, appear to be similar in the data and the model. To validate these observations, we repeated the regressions from Table 3 using the simulated data. The estimation results are shown in Table 9. Again, we translate these estimates into the impact of an additional positive review on each outcome variable for a listing with a 4.5-star rating and ten reviews, which are values close to their averages in the actual data. For ease of comparison, we report the prediction using the estimates from Table 3 based on the actual data in parentheses. For the simulated data, we find that *ceteris paribus* a good review raises the rental

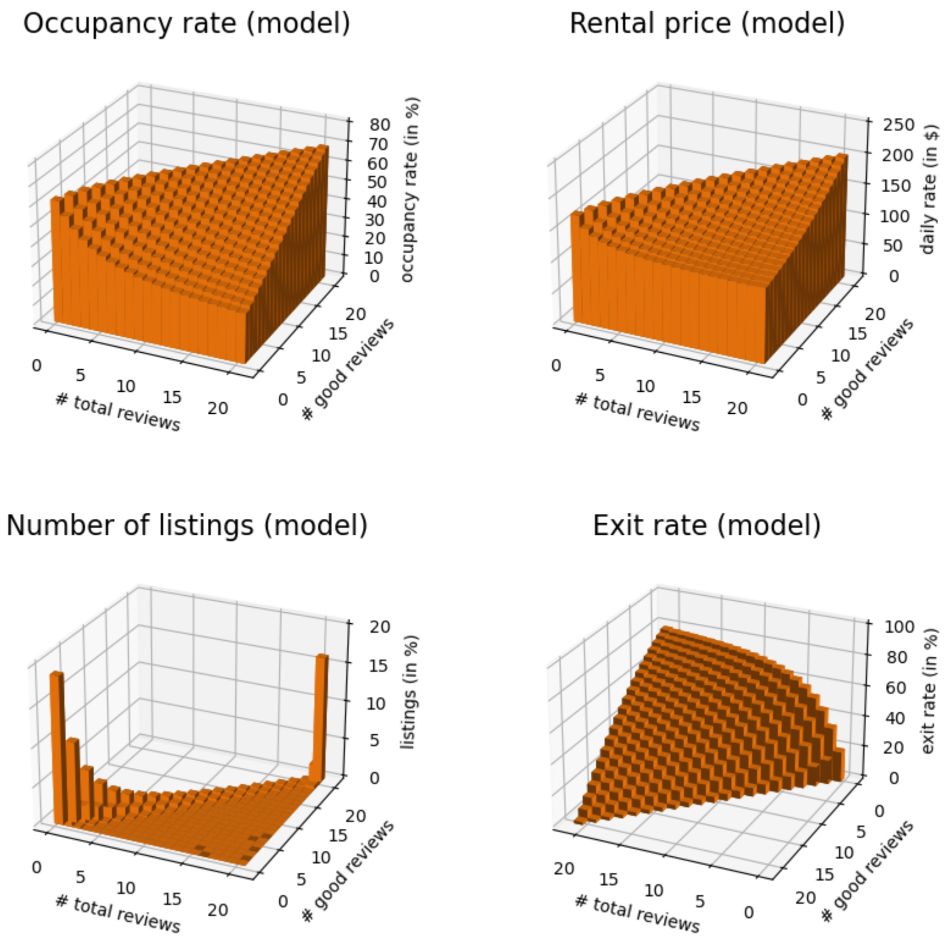

Figure 10: Estimated occupancy rate (top left), rental rate (top right), number of listings (bottom left), and exit rate (bottom right) by number of (good) reviews.

price of by \$1.55 (\$0.98), whereas a bad review decreases the rental price by \$11.79 (\$10.55). Guests generally expect high listing quality and significantly lower their expectations, when being confronted with a negative review. The host must drastically reduce the price to maintain the listing's appeal.

Moreover, Table 3 suggests that a good review increases the occupancy rate of the average host's listing by 0.3 (1.3) percentage points and a bad review decreases it by 8.1 (1.08) percentage points, all else equal. Unlike the hosts in our model, actual hosts might not accurately reflect the rating of their listing in their pricing strategies, as doing so requires solving a complex, dynamic problem. Figure 10 (bottom right) presents the exit rate in the model equilibrium. As expected, state associated with few good reviews feature high exit rates. Hosts of listings with a good rating, on the other hand, are relatively unlikely to leave the market. Directly comparing the model's exit rates with the data is challenging, as our dataset includes exit events for only a few states and for many of these states exit events rarely occur. Our reduced-form estimates suggest that the rating matters more for hosts' exit decisions in the model than in the data. All else equal, the average host is 0.4 (0.2) percentage points less likely to exit the market on any

	Dependent variable:		
	rental rate	occ rate	exit rate
	(1)	(2)	(3)
rental rate		-0.006^{**}	-0.001
		(0.003)	(0.002)
no of reviews	-10.872^{***}	-0.071^{***}	0.071^{***}
	(0.705)	(0.021)	(0.026)
no of good reviews	14.497^{***}	0.095^{***}	-0.095^{***}
	(1.079)	(0.030)	(0.036)
no of reviews ×	-0.105^{***}	-0.001^{***}	0.001^{***}
no of good reviews	(0.012)	(0.000)	(0.000)
Type FE	Yes	Yes	Yes
Observations	59,904	59,904	59,904

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors in parenthesis clustered at type level.

Table 9: Reduced form regression results with simulated data.

given day after receiving a good review; the monthly exit rate increases by 7.8 (0.2) percentage points if the review is bad.

Despite its limitations, our model effectively captures many key aspects of the data. Notably, the model produces a state distribution similar to the empirical one. Additionally, the direction and magnitude of the responses of hosts and guests to reviews in the model are broadly consistent with the actual data. Our counterfactual policies aim to increase the rate at which the quality of new listings is revealed by boosting their occupancy rates. According to our model, there is significant potential to enhance the performance of new listings. Currently, a new listing's occupancy rate is typically 8% below the average, and its rental price is $18 lower than the average. Taken together, new listings generate 17% less revenue than the average listing. Consequently, our model predicts that new listings are around 60% more likely to exit the market in a given month compared to the average listing. In the actual data, new listings are booked 31% less frequently and generate 31% less revenue than the average listing. Therefore, we believe that our counterfactual analysis will yield relatively conservative estimates of the welfare effects of the policies we examine.

9 Counterfactual analysis

Having estimated the model's parameters, we use it to explore how rental prices would need to adjust to ameliorate the cold start problem. Drawing from our theoretical analysis, we are generally interested in the socially optimal price differences among listings. However, solving for 924 welfare-maximizing prices is infeasible. Instead, we consider tax-subsidy schemes that integrate per-booking subsidies or taxes with lump-sum transfers to encourage listings to adjust their prices relative to one another. We allow per-booking taxes or subsidies to vary according to the number of reviews while determining the appropriate lump-sum transfer to ensure that no host experiences a net transfer. This approach effectively compels hosts to modify their prices, with a similar pressure exerted on firms within the same review interval. Since this method offers less flexibility than permitting state-specific price adjustments, the welfare effects of our counterfactual represent a lower bound on the potential welfare gains achievable through altering relative prices.

Equation (18) shows the mean utility of a listing in state x, where the subsidy is denoted by τ. We allow the subsidies to take negative values, in which case we refer to them as taxes.

$$\nu(P(x), x) = \gamma \frac{a + K(x)}{a + b + N(x)} + \beta(x) + \alpha((1 + f)P(x) - \tau(x)) \tag{17}$$

Hence, a subsidy (tax) reduces (increases) the effective price consumers pay and makes a listing more (less) attractive to consumers. To prevent this from merely being a case of transferring money to (from) the host, we deduct (add) the expected subsidy (tax) as a lump-sum transfer from (to) per-period profits:

$$\pi(P(x), x) = q(x)(1 + f)P(x) - T(x) \qquad \text{where} \qquad T(x) = q(x)\tau(x) \tag{18}$$

We search for the subsidy or tax values that maximize long-run equilibrium welfare. Based on our analysis in Section 3, it is unclear whether the cold-start problem exists in this market, but if it does, we expect the optimal tax-subsidy scheme to widen the rental price gap between new and established listings.

9.1 Welfare

Generally, we determine the tax-subsidy scheme that maximizes social welfare by solving the model with varying amounts of subsidies and taxes. We calculate the welfare change as the change in the sum of host and Airbnb revenues, the total operating cost, the total entry cost, and the compensating variation in response to the policy. Since the policy is revenue-neutral, we do not need to separately account for the subsidy cost and tax revenue.

The monthly producer profit is characterized in Equation (19), where all variables take their new oblivious equilibrium values. The first term of Equation (19) captures the hosts' revenues, Airbnb fees, tax revenue, and subsidy cost. The second term represents the operating cost.[24]

$$\sum_{x}^{X} s(x)\left(28q(x)((1+f)P(x)-\tau(x))-((1-\chi(x))\bar{\phi}(x)-\chi(x)\delta T(x)V(x))\right) \tag{19}$$

Equation (20) describes the total entry cost in equilibrium. The entry cost is the number of potential entrants times their type-specific, expected cost of entry. the number of potential type-l entrants equals the total number of listings of that type less the number of type-l listings that are already active.[25]

$$\sum_{l}\left(\frac{J}{4}-\sum_{x}^{X_l} s(x)\right)(\lambda_l\bar{\kappa}_l-(1-\lambda_l)\delta V((0,0,l))) \tag{20}$$

Recall that each listing-week can only be booked once. Ignoring the capacity constraint of listings risks overstating the compensating variation of the subsidy. To address this, we assume that if multiple consumers want to book the same listing, only one makes the booking, while the others must choose the outside option, such as booking a hotel room. As shown by Williams (1977) and Small and Rosen (1981), if the random taste shocks are independently and identically Gumbel distributed and the utility is linear in income, the expected utility of a single consumer is the natural logarithm of the sum of the mean utilities, $\ln(1+\sum_x^X s(x)\nu(P(x),x))$, plus a constant of integration. If capacities were not constrained, consumer surplus would be the total market size 28μ times the expected utility. However, since capacities are constrained, not every guest can book their preferred listing. From Equation (5), we know that under the capacity constraint, in expectation, $q(x)$ consumers book a particular listing in state x. Conversely, the expected number of guests who want to book that listing is $-\ln(1-q(x))$, which is greater than $q(x)$. When calculating consumer surplus, we correct for this difference, as shown in Equation (21).

[24] For a derivation of the expected operating cost, see Online Appendix B.4.
[25] For a derivation of the expected entry cost see Online Appendix B.5.

$$28\left(\mu-\sum_x^X s(x)(-\ln(1-q(x))-q(x))\right)\ln\left(1+\sum_x^X s(x)\nu(P(x),x)\right)+\text{constant} \qquad (21)$$

We calculate the compensating variation as the change in consumer surplus divided by the price coefficient α (McFadden, 2012; Small and Rosen, 1981).

9.2 Changing the price difference between new and established listings

In this section, we focus solely on the welfare gain that can be achieved by diverting demand through changing relative prices between listings depending on their number of reviews. To this end, we determine review intervals such that taxes or subsidies vary between intervals, but are uniform within intervals. Specifically, we define intervals in steps of five reviews, except for the the first five review, which we split into smaller intervals reflecting that the value of information decreases rapidly for the first reviews. Hence, intervals are $[0-1, 2-5, 6-10, 11-15, 16-20]$. Moreover, we require that taxes and subsidies balance out in the aggregate to avoid changes in price level driving the welfare effects. As we have discussed in Section 3, a social planner may choose a different price level from the one arising in the Nash equilibrium to alleviate the problems of aggregate entry or exit inefficiency and market power which are both orthogonal to the reallocation of demand among existing listings. Note that this implies that we simplified the problem from solving for 924 optimal prices to solving for four optimal taxes or subsidies.

Our findings are summarized in Table 10 and Figure 11. We discover that the optimal tax-subsidy scheme leads to a reduction in prices for listings with five or fewer reviews, while the prices for listings with more than five reviews increase; the overall change in the average rental price is minimal. The most significant price decrease, exceeding 16%, occurs for listings with no or one review, while listings with more than 15 reviews see the highest price increase of nearly 10%. Additionally, the price difference between listings with more than 15 reviews and those with fewer than two reviews rises by more than $47, an increase by more than 130%.

These price adjustments lead to changes in occupancy rates which increase for lesser-reviewed listings and decrease for often-reviewed ones. Notably, listings with no or one review experience an occupancy rate increase of almost 14%, enhancing their chances of receiving additional reviews. This contributes to their reduced numbers in the new equilibrium, but it is not the sole factor. The decline in entrant listings results not only from faster learning but also from lower per-period profits as a result of lower prices, which decreases both entry rates and market retention, as shown in the bottom right panel of Figure 11.

Conversely, for listings with more than 15 reviews, the price increase appears to enhance profitability and reduce exit rates, leading to an increase in their numbers. One possible explanation is that a mandatory price increase could enable these incumbents to achieve a more collusive market outcome. As a results consumers benefit from a wider array of established, high-quality options, even though their prices are higher. The higher profitability associated with being an established listing may further incentivize new entrants to ”climb the ladder,” resulting in a decline in their numbers, while listings with at least 16 reviews increase by over

50%.

Overall, this policy leads to a welfare increase of $338,518 per month. In the status-quo equilibrium, the monthly total host revenue on Airbnb in Manhattan amounts to almost $4 million, hence the gain correspond to roughly 8.5% of that revenue. Consumers are the main beneficiaries of the tax-subsidy scheme, gaining more than $350,000 in surplus. In contrast, hosts experience a significant reduction in their surplus, dropping by over 11%. Airbnb benefits from an 8% increase in revenue.

	total	*listings within review interval*				
		[0-1]	[2-5]	[6-10]	[11-15]	[16-20]
Δ per-period welfare	$338,518	–	–	–	–	–
	(6.5%)	–	–	–	–	–
Δ per-period consumer surplus	$350,846	–	–	–	–	–
	(8.4%)	–	–	–	–	–
Δ per-period host surplus	-$57,309	–	–	–	–	–
	(-11.1%)	–	–	–	–	–
Δ per-period Airbnb revenue	$44,981	–	–	–	–	–
	(8.0%)	–	–	–	–	–
Δ average rental price	$2.8	-$27.6	-$7.0	$4.6	$10.0	$19.6
	(1.5%)	(-16.6%)	(-3.9%)	(2.4%)	(5.0%)	(9.7%)
Δ average occupancy rate	1.6 ppt	8.3 ppt	2.2 ppt	-1.1 ppt	-2.5 ppt	-4.4 ppt
	(2.4%)	(13.9%)	(3.4%)	(-1.6%)	(-3.6%)	(-6.4%)
Δ # listings	53	-66	-28	-2	4	144
	(4.6%)	(-18.1%)	(-9.4%)	(-1.4%)	(5.6%)	(54%)

Note: Percentage changes are indicated relative to prices, number of listings, or occupancy in the status quo.

Table 10: Effects of the socially-optimal price gap between entrants and incumbents.

Consequently, the optimal tax-subsidy scheme encourages some guests to book and review listings that have very few reviews rather than listings with many reviews. This increases the availability of information regarding product quality for all consumers and aids in identifying more high-quality listings. At the same time, changes in prices and booking probabilities also impact per-period profits, which subsequently affect entry and exit dynamics. If there is heterogeneity in firm value for consumers, this will be reflected in the optimal tax-subsidy scheme as well; the social planner would want to avoid exit by highly valuable market participants. In our setting, consumers prioritize preventing the exit of established incumbent listings over relatively newer listings, which are generally perceived to be of lower quality. This motive reinforces the benefits of increasing the price gap between entrants and incumbents: The welfare loss from the increase in exit rates of new listings is more than compensated by the gain from fewer exits of established listings.

It is important to note that this aspect is not specific to our approach of redirecting demand through prices. Any intervention aimed at addressing the social learning externality will affect profits and has to take this into account. For instance, if we implemented a change in ratings design to shift demand from incumbents to entrants as in Vellodi (2022), this would similarly not only affect the speed of social learning, but also trigger price and profits changes with immediate consequences for entry and exit. This is a channel which is lacking in most theoretical papers on the cold-start problem.

Still, one might question how much of the welfare effect stems from changes in host profits versus the increase in learning speed alone. When we calculate welfare under the prices resulting from the optimal tax-subsidy scheme, while freezing revenues at their zero-tax level, such that entry and exit are influenced solely by changes in transition probabilities, the total welfare effect amounts to $253,810. Consumer surplus, host surplus, and Airbnb revenue changes follow a similar pattern as in Table 10. Hence, the change in profits attributed to roughly 25% of the welfare gain and plays a significant role when allowing for endogenous entry and exit. Prior to implementing any intervention, how it may affect profits must be carefully considered.

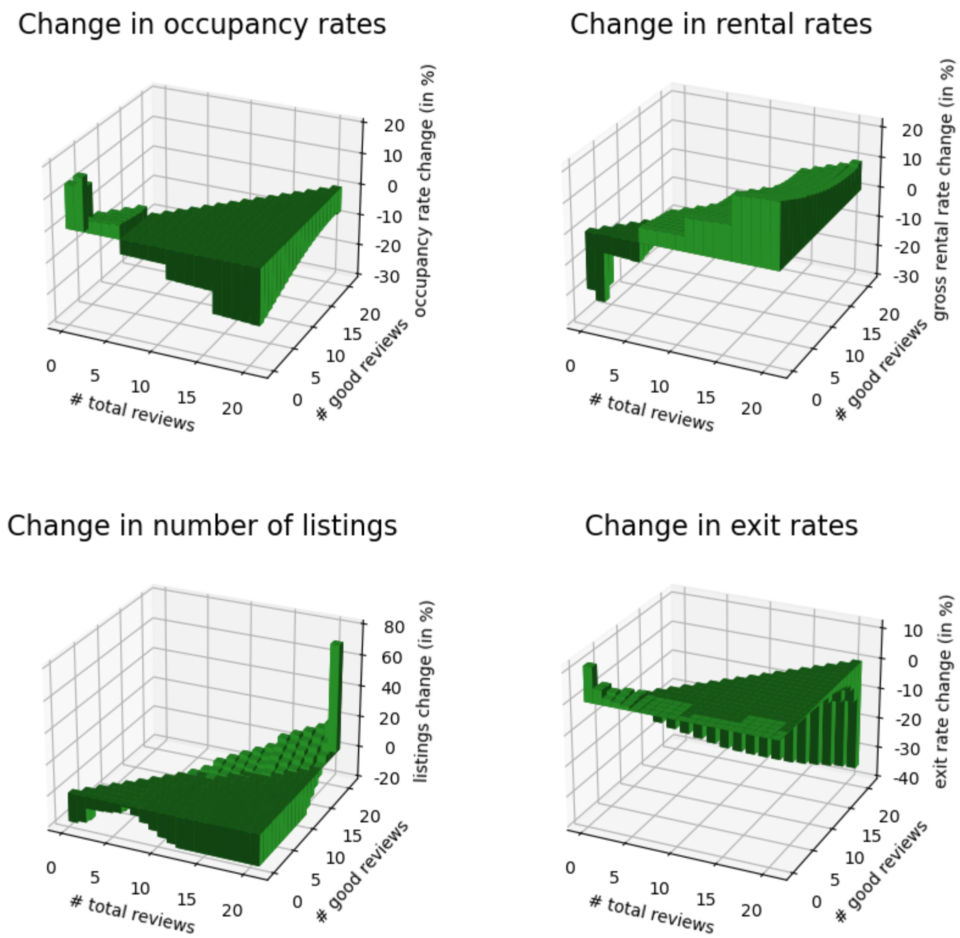

Figure 11: Counterfactual changes in occupancy rate (top left), rental rate (top right), number of listings (bottom left), and exit rate (bottom right) by number of (good) reviews.

9.3 Setting optimal prices for listings with no reviews

As mentioned in the introduction, Airbnb appears to recognize the cold-start problem, as it recommends that new listings provide a price discount for their first guests in order to "get their initial booking faster."[26] Consequently, in this counterfactual analysis, we examine the socially optimal prices for these brand-new listings with no reviews and assess the total welfare gains Airbnb could realize by encouraging new hosts to fully comply with this pricing suggestion. Given that there are four types of brand-new listings, we optimize over four distinct prices. At the same time, we allow the remaining hosts to respond to price changes for new listings, such that our analysis takes into account the full response of the market.

Our results are summarized in Table 11. Also in this counterfactual, we find that entrant prices decrease, with larger relative decreases for lower-type listings. However, in contrast to the previous counterfactual, the total number of listings decrease as well. This is because lowering entrant prices decreases profits, and thus, their entry incentives, despite their chances of receiving a review and making higher profits in the future increasing. Moreover, decreasing entrant prices increases competition on Airbnb and decreases profits for incumbents as well, if only slightly. This leads to a decrease in host surplus as well as Airbnb revenue by around 2%, while consumers are the beneficiaries of this policy, with a welfare increase by around 1%. Overall welfare increases by only $20,646, or 0.5% of total host revenue in the status quo.

	total	*listings with no review of type*			
		1	2	3	4
Δ per-period welfare	$20,646	–	–	–	–
	(0.4%)	–	–	–	–
Δ per-period consumer surplus	$43,507	–	–	–	–
	(1.0%)	–	–	–	–
Δ per-period host surplus	-$10,712	–	–	–	–
	(-2.1%)	–	–	–	–
Δ per-period Airbnb revenue	-$12,149	–	–	–	–
	(-2.2%)	–	–	–	–
Δ average rental price	-$3.4	-$19.3	-$19.5	-$18.9	-$17.4
	(-1.8%)	(-14.5%)	(-12.1%)	(-10.5%)	(-8.2%)
Δ average occupancy rate	1.2 ppt	6.4 ppt	7.0 ppt	6.6 ppt	5.6 ppt
	(1.7%)	(14.3%)	(11.5%)	(9.8%)	(7.5%)
Δ # listings	-21	-6	-6	-4	-2
	(-1.9%)	(-9.1%)	(-8.4%)	(-7.6%)	(6.2%)

Note: Percentage changes are indicated relative to prices, number of listings, or occupancy in the status quo.

Table 11: Effects of socially-optimal prices for listings with no reviews only.

[26] See https://www.airbnb.ca/resources/hosting-homes/a/how-to-set-a-pricing-strategy-15

When considering the impact of incentivizing social learning on profits and seller participation in the market, there can be a trade-off between accelerating social learning and maintaining entry levels or preventing the exit of high-value products. In this counterfactual scenario, while the benefits of enhancing the speed of social learning remain, it becomes evident that the social planner may be reluctant to compromise market participation of higher-type listings, and chooses to distort prices less, by only 8% compared to 15% for low-type listings. Furthermore, the welfare gains that can be solely obtained from reducing prices for entrants are relatively modest.

It is also noteworthy that in this counterfactual, although consumer surplus rises, Airbnb's revenue declines. This occurs because both market size and prices decrease, yet the benefits derived from more information on product quality more than offset these losses for consumers. Therefore, we should not expect that Airbnb's incentives are necessarily aligned with those of the social planner.

9.4 Discussion

By implementing a tax-subsidy scheme, we have demonstrated that significant welfare gains can be achieved by addressing the cold start problem in the Airbnb market. However, it is important to note that we do not advocate for the implementation of this policy. Market conditions have evolved since the period of our study from 2016 to 2019. As of September 5th, 2023, the City of New York has mandated that short-term rentals must register with the Mayor's Office of Special Enforcement.[27] This measure aims to curb short-term home-sharing and support the long-term rental market.

Calder-Wang (2021) finds that although Airbnb's presence generally contributes positively to welfare, it has the adverse consequence of increasing rents in New York City, thereby making most households worse off. Our model does not account for housing market effects and they are not covered by our analysis.

The primary objective of our study is to illustrate how the inefficiently slow speed of learning affects the market and analyze the trade-off involved in interventions aimed at remedying the problem when taking the supply side into account. The tax policy we analyze serves as an auxiliary tool to highlight these effects. The general nature of the problem suggests that virtually any digital market with a review system may suffer from significant welfare losses due to the too-slow accumulation of reviews, but addressing the cold-start problem involves understanding the consequences for profits and market participation.

10 Conclusion

In this paper, we aim to analyze the cold-start problem and its interactions with endgenous supply-side decisions, such as pricing, entry and exit. Specifically, we demonstrate that the cold-start problem affects Airbnb and that its impact is substantial.

[27] See https://www.nyc.gov/site/specialenforcement/registration-law/registration.page.

To achieve this, we have developed a model in which Airbnb guests learn from each other's experiences through reviews, while hosts use these reviews to inform their pricing, entry, and exit decisions. Our findings indicate that reviews significantly influence guests' booking decisions and host profits.

We present counterfactual results aligned with the cold-start problem, showing that long-term social welfare can be improved by reducing the price of less-reviewed listings. This strategy boosts demand and accelerates the learning process.

Specifically, we find that implementing a tax-subsidy scheme that more than doubles the price difference between listings with more than 15 and less than 2 reviews leads to a welfare gain equivalent to roughly 8.5% of Airbnb host revenue. Our estimates suggest that 75% of this welfare gain is attributed to the accelerated learning process alone. Additionally, we evaluate a policy that requires low prices for unreviewed listings only and find that while it generates some improvements, the gains are relatively modest.

We conclude that the cold-start problem represents a significant challenge, and addressing it could lead to substantial gains for Airbnb and potentially other digital platforms. We recommend that this issue be considered in the design of platforms and regulations of markets that feature social learning more broadly.

Appendix

A.1 Proof of Lemma 1

In $t = 1$, firm j maximizes

$$p_{j1}q_{j1} + v_r\delta q_{E1}\omega^e_{E1}\pi^g_{j2} + v_r\delta q_{E1}(1-\omega^e_{E1})\pi^b_{j2} + (1 - v_r\delta q_{E1})\pi^0_{j2}.$$

The necessary and sufficient conditions for a maximum are as follows.

$$q_{j1} - p^*_{j1}q_{j1}(1-q_{j1}) + v_r\delta q_{E1}(1-q_{E1})\frac{\partial(-p_{E1}+p_{I1})}{\partial p_{j1}} \sum_{i\in\{0,g,b\}} \phi_i\pi^i_{j2} = 0 \quad \text{(FOC)}$$

$$q_{j1}(1-q_{j1})(p^*_{j1}(1-q_{j1})-1) + v_r\delta q_{E1}(1-q_{E1})^2\left(\frac{\partial(-p_{E1}+p_{I1})}{\partial p_{j1}}\right)^2 \sum_{i\in\{0,g,b\}} \phi_i\pi^i_{j2} < 0 \quad \text{(SOC)}$$

Rearranging the FOC yields p^*_{j1}.

$$p^*_{j1} = \frac{1}{1-q^*_{j1}} + \frac{\partial(-p_{E1}+p_{I1})}{\partial p_{j1}} \sum_{i\in\{0,g,b\}} v_r\delta\phi_i\pi^i_{j2} \tag{22}$$

Substituting p^*_{j1} into the SOC reveals that the sufficient condition for a maximum is satisfied.

In $t = 2$, j maximizes $p^i_{j2}q^i_{j2}$. It is straightforward to verify that $p^*_{j2} = 1/(1-q_{j2})$. Hence, $\pi^i_{E2} = q^i_{E2}/(1-q^i_{E2})$ which is increasing and convex in ω^e_{E2}.

$$\frac{\partial}{\partial\omega^e_{E2}}\left(\frac{q_{E2}}{1-q_{E2}}\right) = \frac{q_{E2}(1-q_{E2})^2 + q^2_{E2}(1-q_{E2})}{(1-q_{E2})^2} = \frac{q_{E2}}{1-q_{E2}} > 0$$

$$\frac{\partial^2}{\partial(\omega^e_{E2})^2}\left(\frac{q_{E2}}{1-q_{E2}}\right) = \frac{q_{E2}}{1-q_{E2}} > 0$$

By Jensen's inequality $\sum_{i\in\{0,g,b\}} \phi_i\pi^i_{E2}$ is larger than zero.

$$\begin{aligned}\omega^e_{E1}\pi_{E2}\left(\frac{a+1}{a+b+1}\right) + (1-\omega^e_{E1})\pi_{E2}\left(\frac{a}{a+b+1}\right) - \pi_{E2}\left(\omega^e_{E1}\right)\\ > \pi_{E2}\left(\omega^e_{E1}\frac{a+1}{a+b+1} + (1-\omega^e_{E1})\frac{a}{a+b+1}\right) - \pi_{E2}(\omega^e_{E1}) = 0\end{aligned}$$

Notice that $\partial p^*_{E1}/\partial v_r < 0$ and $\partial p^*_{E1}/\partial\delta < 0$. As $q_{It} = 1 - q_{Et}$, $\pi^i_{I2} = (\pi^i_{E2})^{-1}$ and π_{I2} is decreasing and convex in ω^e_{E2}. Again, by Jensen's inequality $\sum_{i\in\{0,g,b\}} \phi_i\pi^i_{I2}$ is larger than zero and $\partial p^*_{I1}/\partial v_r > 0$, as well as $\partial p^*_{I1}/\partial\delta > 0$.

A.2 Proof of Proposition 1

Part (i) The social planner solves the following maximization problem.

$$\max_{\tilde{p}_1} \left(cs_1 + \pi_{E1} + \pi_{I1} + \mathbb{E}_2[cs_2 + \pi_{E2} + \pi_{I2} | p_{E1}, p_{I1}]\right)$$

We write $\pi_{E1} + \pi_{I1}$ as $p_{I1} + q_{E1}\tilde{p}_1$. For brevity, denote $\ln\left(1 + \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)\right)$ by u_2, where $\tilde{\omega}_2$ and $\tilde{p}_2$ are the difference in second-stage quality expectations and Nash-equilibrium prices, respectively.

Before proceeding with the FOC, we establish two helpful facts:

1. u_2 is increasing and convex in w_{E2}:

$$\frac{\partial u_2}{\partial w_{2E}} = \frac{1}{\left(1 + \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)\right)} \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right) = q_{E2} > 0$$
$$\frac{\partial^2 u_2}{\partial w_{2E}^2} = q_{E2}(1 - q_{E2}) > 0$$

2. $\pi_{E2} + \pi_{I2}$ is convex in w_{E2}:

$$\pi_{E2} + \pi_{I2} = \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right) + \frac{1}{\exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)}$$
$$\frac{\partial(\pi_{E2} + \pi_{I2})}{\partial w_{2E}} = \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right) - \frac{1}{\exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)}$$
$$\frac{\partial^2(\pi_{E2} + \pi_{I2})}{\partial w_{2E}^2} = \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right) + \frac{1}{\exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)} > 0$$

The necessary and sufficient conditions for a maximum are as follows.

$$-q_{E1}(1 - q_{E1})\tilde{p}_1^s + q_{E1} - q_{E1} - q_{E1}(1 - q_{E1}) \sum_{i \in \{0,g,b\}} v_r \delta \phi_i \left(u_2^i + \pi_{E2}^i + \pi_{I2}^i\right) = 0 \quad \text{(FOC)}$$

$$-q_{E1}(1 - q_{E1}) + q_{E1}\left(1 - q_{E1}\right)\left(2q_{E1} - 1\right)\tilde{p}_1^s - q\left(1 - q_{E1}\right)\left(2q_{E1} - 1\right) \sum_{i \in \{0,g,b\}} v_r \delta \phi_i \left(u_2^i + \pi_{E2}^i + \pi_{I2}^i\right) < 0 \quad \text{(SOC)}$$

Rearranging the FOC yields $\tilde{p}_1^s$.

$$\tilde{p}_1^s = -\sum_{i \in \{0,g,b\}} v_r \delta \phi_i \left(u_2^i + \pi_{E2}^i + \pi_{I2}^i\right) \tag{23}$$

From facts 1. and 2. and Jensen's inequality, it follows that $\tilde{p}_1^s < 0$. It easy to see that the SOC is satisfied at the socially optimal price difference, i

Part (ii) From Equation (22) and Equation (23), we know:

$$\tilde{p}_1^s < \tilde{p}_1^* \iff -\sum_i v_r \delta \phi_i u_2^i < \frac{1}{1 - q_{E1}^*(\tilde{\omega}_1)} - \frac{1}{q_{E1}^*(\tilde{\omega}_1)}$$

$$= \exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)) - \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1))} \tag{24}$$

We will show this part of the proposition in two steps: First, we show that the above inequality is satisfied at $\tilde{\omega}_1 = 0$. Second we show that, as $\tilde{\omega}_1$ increases, the increase in $\tilde{p}_1^s$ is smaller than the increase in $\tilde{p}_1^*$ for all $\tilde{\omega}_1$, as long as $\sum_i \delta v_r \phi_i(\pi_{I2}^i - \pi_{E2}^i) < 1$.

Step 1: At $\tilde{\omega}_1 = 0, \tilde{p}_1^s < \tilde{p}_1^*$: $\tilde{p}_1^*$ is given implicitly by the difference in Equation (22) for entrant and incumbent:

$$\tilde{p}_1^*(\tilde{\omega}_1) = \exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)) - \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1))} - \sum_i \delta v_r \phi_i(\pi_{E2}^i + \pi_{I2}^i) \tag{25}$$

Notice that for any ω_{E1}, Equation (24) implies that, if $\tilde{\omega}_1 = 0$, $\tilde{p}_1^* < 0$. Therefore,

$$\exp(-\tilde{p}_1^*(0)) - \frac{1}{\exp(-\tilde{p}_1^*(0))} > 0 > -\sum_i v_r \delta \phi_i u_2^i.$$

Step 2: $\frac{d\tilde{p}_1^s}{d\tilde{\omega}_1} < \frac{d\tilde{p}_1^*}{d\tilde{\omega}_1}$ iff $\sum_i \delta v_r \phi_i(\pi_{I2}^i - \pi_{E2}^i) < 1$: Since $\frac{\partial \tilde{\omega}_1}{\partial \omega_{I1}} = -1$, we consider a marginal decrease in ω_{I1} here:

$$\frac{d(-\sum_i v_r \delta \phi_i u_2^i)}{d\omega_{I1}} = \sum_i v_r \delta \phi_i q_{E2}^{i*} \left(1 + \frac{d\tilde{p}_2^*}{d\omega_{I2}}\right) > 0 \text{ since } -1 < \frac{d\tilde{p}_2^*}{d\omega_{I2}} < 0.$$

Hence, the LHS of inequality (24) is decreasing as ω_{I1} decreases or $\tilde{\omega}_1$ increases.

$$\frac{d\left(\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)) - \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1))}\right)}{d\omega_{I1}} = -\left(\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)) + \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1))}\right)\left(1 + \frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}}\right)$$

As long as $-1 < \frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}}$, the RHS of inequality (24) increases as ω_{I1} decreases or $\tilde{\omega}_1$ increases.

Using the Implicit Function Theorem on Equation (25), we can derive $\frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}}$:

$$\frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}} = -\frac{\exp(\tilde{\omega}_1 - \tilde{p}_1(\tilde{\omega}_1)) + \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1(\tilde{\omega}_1))} + \sum_i \delta v_r \phi_i(\pi_{I2}^i - \pi_{E2}^i)}{\exp(\tilde{\omega}_2 - \tilde{p}_1(\tilde{\omega}_1)) + \frac{1}{\exp(\tilde{\omega}_1 - \tilde{p}_1(\tilde{\omega}_1))} + 1}$$

Hence, $\frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}} > -1$ iff $\sum_i \delta v_r \phi_i(\pi_{I2}^i - \pi_{E2}^i) < 1$.

References

Aguirregabiria, Victor, Allan Collard-Wexler, and Stephen P Ryan, "Dynamic games in empirical industrial organization," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 225–343.

Almagro, Milena and Tomás Domınguez-Iino, "Location sorting and endogenous amenities: Evidence from Amsterdam," *Working Paper*, 2024.

Anderson, Michael and Jeremy Magruder, "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database," *The Economic Journal*, 2012, *122* (563), 957–989.

Bao, Ying, Limin Fang, and Matthew Osborne, "The Effect of Quality Disclosure on Firm Entry and Exit Dynamics: Evidence from Online Review Platforms," *Available at SSRN 4804928*, 2024.

Barwick, Panle Jia and Parag A Pathak, "The costs of free entry: an empirical study of real estate agents in Greater Boston," *The RAND Journal of Economics*, 2015, *46* (1), 103–145.

Bergemann, Dirk and Juuso Välimäki, "Learning and strategic pricing," *Econometrica*, 1996, *64* (5), 1125–1149.

_ and _ , "Market diffusion with two-sided learning," *The RAND Journal of Economics*, 1997, pp. 773–795.

_ and _ , "Experimentation in markets," *The Review of Economic Studies*, 2000, *67* (2), 213–234.

Besanko, David, Ulrich Doraszelski, and Yaroslav Kryukov, "The economics of predation: What drives pricing when there is learning-by-doing?," *American Economic Review*, 2014, *104* (3), 868–97.

_ , _ , and _ , "How efficient is dynamic competition? the case of price as investment," *American Economic Review*, 2019, *109* (9), 3339–64.

_ , _ , _ , and Mark Satterthwaite, "Learning-by-doing, organizational forgetting, and industry dynamics," *Econometrica*, 2010, *78* (2), 453–508.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou, "Geography, transportation, and endogenous trade costs," *Econometrica*, 2020, *88* (2), 657–691.

Cabral, Luis and Ali Hortacsu, "The dynamics of seller reputation: Evidence from eBay," *The Journal of Industrial Economics*, 2010, *58* (1), 54–78.

_ and Michael H Riordan, "The learning curve, market dominance, and predatory pricing," *Econometrica*, 1994, *62* (5), 1115–1140.

_ and _ , "The learning curve, predation, antitrust, and welfare," *The Journal of Industrial Economics*, 1997, *45* (2), 155–169.

Calder-Wang, Sophie, "The distributional impact of the sharing economy on the housing market," *Working Paper*, 2021.

Che, Yeon-Koo and Johannes Hörner, "Recommender Systems as Mechanisms for Social Learning," *The Quarterly Journal of Economics*, 2018, *133* (2), 871–925.

Chen, Jiawei, "How do switching costs affect market concentration and prices in network industries?," *The Journal of Industrial Economics*, 2016, *64* (2), 226–254.

Chen, Yanyou and Daniel Yi Xu, "A Structural Empirical Model of R&D, Firm Heterogeneity, and Industry Evolution," *The Journal of Industrial Economics*, 2023, *71* (2), 323–353.

Chevalier, Judith A and Dina Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, 2006, *43* (3), 345–354.

Ching, Andrew T, "A dynamic oligopoly structural model for the prescription drug market after patent expiration," *International Economic Review*, 2010, *51* (4), 1175–1207.

Collard-Wexler, Allan, "Demand fluctuations in the ready-mix concrete industry," *Econometrica*, 2013, *81* (3), 1003–1037.

Dickstein, Michael J, "Efficient provision of experience goods: Evidence from antidepressant choice," *Working Paper*, 2021.

Doraszelski, Ulrich and Kenneth L Judd, "Avoiding the curse of dimensionality in dynamic stochastic games," *Quantitative Economics*, 2012, *3* (1), 53–93.

_ and _ , "Dynamic stochastic games with random moves," *Quantitative Marketing and Economics*, 2019, *17* (1), 59–79.

_ and Mark Satterthwaite, "Computable Markov-perfect industry dynamics," *The RAND Journal of Economics*, 2010, *41* (2), 215–243.

Dubé, Jean-Pierre H, Günter J Hitsch, and Pradeep K Chintagunta, "Tipping and concentration in markets with indirect network effects," *Marketing Science*, 2010, *29* (2), 216–249.

Ericson, Richard and Ariel Pakes, "Markov-perfect industry dynamics: A framework for empirical work," *The Review of Economic Studies*, 1995, *62* (1), 53–82.

Fang, Limin, "The effects of online review platforms on restaurant revenue, consumer learning, and welfare," *Management Science*, 2022, *68* (11), 8116–8143.

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

Ferrari, Silvia and Francisco Cribari-Neto, "Beta regression for modelling rates and proportions," *Journal of Applied Statistics*, 2004, *31* (7), 799–815.

Fershtman, Chaim and Ariel Pakes, "Dynamic games with asymmetric information: A framework for empirical work," *The Quarterly Journal of Economics*, 2012, *127* (4), 1611–1661.

Fradkin, Andrey and David Holtz, "Do incentives to review help the market? Evidence from a field experiment on Airbnb," *American Economic Review: Insights*, 2022.

_ , Elena Grewal, and David Holtz, "Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on Airbnb," *Marketing Science*, 2021, *40* (6), 1013–1029.

Frechette, Guillaume R, Alessandro Lizzeri, and Tobias Salz, "Frictions in a competitive, regulated market: Evidence from taxis," *American Economic Review*, 2019, *109* (8), 2954–92.

Hagiu, Andrei and Julian Wright, "Platforms and the exploration of new products," *Management Science*, 2020, *66* (4), 1527–1543.

Hortaçsu, Ali, Olivia R Natan, Hayden Parsley, Timothy Schwieg, and Kevin R Williams, "Organizational Structure and Pricing: Evidence from a Large U.S. Airline," *The Quarterly Journal of Economics*, 09 2023, *139* (2), 1149–1199.

Huang, Yufeng, "Pricing Frictions and Platform Remedies: The Case of Airbnb," *Available at SSRN 3767103*, 2022.

Igami, Mitsuru and Kosuke Uetake, "Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2016," *The Review of Economic Studies*, 2020, *87* (6), 2672–2702.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, 1982, *50* (3), 649–670.

Kellogg, Ryan, "The effect of uncertainty on investment: Evidence from Texas oil drilling," *American Economic Review*, 2014, *104* (6), 1698–1734.

Kremer, Ilan, Yishay Mansour, and Motty Perry, "Implementing the "wisdom of the crowd"," *Journal of Political Economy*, 2014, *122* (5), 988–1012.

Lewis, Gregory and Georgios Zervas, "The welfare impact of consumer reviews: A case study of the hotel industry," *Working Paper*, 2016.

Luca, Michael, "Reviews, reputation, and revenue: The case of Yelp. com," *Working Paper*, 2016, (12-016).

McFadden, Daniel, "Computing willingness–to–pay in random utility models," in "Trade, theory and econometrics" 2012, pp. 275–296.

Pallais, Amanda, "Inefficient hiring in entry-level labor markets," *American Economic Review*, 2014, *104* (11), 3565–3599.

Proserpio, Davide, Wendy Xu, and Georgios Zervas, "You get what you give: theory and evidence of reciprocity in the sharing economy," *Quantitative Marketing and Economics*, 2018, *16*, 371–407.

Reimers, Imke and Joel Waldfogel, "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review*, 2021, *111* (6), 1944–71.

Rossi, Michelangelo, "Competition and Reputation in an Online Marketplace: Evidence from Airbnb," *Management Science*, 2024, *70* (3), 1357–1373.

Small, Kenneth A and Harvey S Rosen, "Applied welfare economics with discrete choice models," *Econometrica*, 1981, pp. 105–130.

Sweeting, Andrew, Dun Jia, Shen Hui, and Xinlu Yao, "Dynamic price competition, learning-by-doing, and strategic buyers," *American Economic Review*, 2022, *112* (4), 1311–1333.

_, James W Roberts, and Chris Gedge, "A model of dynamic limit pricing with an application to the airline industry," *Journal of Political Economy*, 2020, *128* (3), 1148–1193.

Takahashi, Yuya, "Estimating a war of attrition: The case of the us movie theater industry," *American Economic Review*, 2015, *105* (7), 2204–2241.

Vellodi, Nikhil, "Ratings design and barriers to entry," *Working Paper*, 2022.

Weintraub, Gabriel Y, C Lanier Benkard, and Benjamin Van Roy, "Markov perfect industry dynamics with many firms," *Econometrica*, 2008, *76* (6), 1375–1411.

_ , _ , and _ , "Computational methods for oblivious equilibrium," *Operations Research*, 2010, *58* (4-part-2), 1247–1265.

Williams, Huw CWL, "On the formation of travel demand models and economic evaluation measures of user benefit," *Environment and Planning A*, 1977, *9* (3), 285–344.

Williams, Kevin R, "The welfare effects of dynamic pricing: Evidence from airline markets," *Econometrica*, 2022, *90* (2), 831–858.

Wollmann, Thomas G, "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.

Wu, Chunhua, Hai Che, Tat Y Chan, and Xianghua Lu, "The economic value of online reviews," *Marketing Science*, 2015, *34* (5), 739–754.

Zervas, Georgios, Davide Proserpio, and John W Byers, "A first look at online reputation on Airbnb, where every stay is above average," *Marketing Letters*, 2021, *32*, 1–16.

Supplementary Online Appendix

B.1 Transition matrix

Conditional on $\mathbf{P}_t, \mathbf{s}_t$, and l, transition matrix $\mathbf{T}$ is given by the following matrix.

$(K_t, N_t)\backslash(K_{t+1}, N_{t+1})$	$(0,0)$	$(0,1)$	$(1,1)$	$(0,2)$	$(1,2)$	$(2,2)$	...	$(\bar{N},\bar{N})$
$(0,0)$	$\rho^0(0,0)$	$\rho^b(0,0)$	$\rho^g(0,0)$	0	0	0	...	0
$(0,1)$	0	$\rho^0(0,1)$	0	$\rho^b(0,1)$	$\rho^g(0,1)$	0	...	0
$(1,1)$	0	0	$\rho^0(1,1)$	0	$\rho^b(1,1)$	$\rho^g(1,1)$	...	0
$(0,2)$	0	0	0	$\rho^0(0,2)$	0	0	...	0
$(1,2)$	0	0	0	0	$\rho^0(1,2)$	0	...	0
$(2,2)$	0	0	0	0	0	$\rho^0(2,2)$	...	0
...	...	...	...	...	...	...	...	
$(\bar{N},\bar{N})$	0	0	0	0	0	0	...	1

Conditional on $\mathbf{P}_t, \mathbf{s}_t$, and l expanded transition matrix $\mathbf{F}$ is given by the following matrix.

$(K_t, N_t)\backslash(K_{t+1}, N_{t+1})$	$(0,0)$	$(0,1)$	$(1,1)$	$(0,2)$	$(1,2)$	$(2,2)$	...	$(\bar{N},\bar{N})$	*inactive*
$(0,0)$	$(1-\chi(0,0))\rho^0(0,0)$	$(1-\chi(0,0))\rho^b(0,0)$	$(1-\chi(0)\rho^g(0,0)$	0	0	0	...	0	$\chi(0,0)$
$(0,1)$	0	$(1-\chi(0,1))\rho^0(0,1)$	0	$(1-\chi(0,1))\rho^b(0,1)$	$(1-\chi(0,1))\rho^g(0,1)$	0	...	0	$\chi(0,1)$
$(1,1)$	0	0	$(1-\chi(1,1))\rho^0(1,1)$	0	$(1-\chi(1,1))\rho^b(1,1)$	$(1-\chi(1,1))\rho^g(1,1)$	...	0	$\chi(1,1)$
$(0,2)$	0	0	0	$(1-\chi(0,2))\rho^0(0,2)$	0	0	...	0	$\chi(0,2)$
$(1,2)$	0	0	0	0	$(1-\chi(1,2))\rho^0(1,2)$	0	...	0	$\chi(1,2)$
$(2,2)$	0	0	0	0	0	$(1-\chi(0,2))\rho^0(0,2)$	...	0	$\chi(0,2)$
...	...	...	...	...	...	...	...		...
$(\bar{N},\bar{N})$	0	0	0	0	0	0	...	$1-\chi(\bar{N},\bar{N})$	$\chi(\bar{N},\bar{N})$
inactive	λ	0	0	0	0	0	...	0	$1-\lambda$

B.2 First- and second-order derivatives of the value function

$$
\begin{aligned}
v'(x) =&28(q_0(P_0(x), x) + q_0'(p_t, x)P_0(x)) + (1 - \chi_0(P_0(x), x))\delta T_0'(P_0(x), x)V_0(x) \\
v''(x) =&28(2q_0'(p_t, x) + q_0''(P_0(x), x)P_0(x)) + (1 - \chi_0(P_0(x), x))\delta T_0''(P_0(x), x)V_0(x) - \chi_0'(P_0(x), x)\delta T_0'(P_0(x), x)
\end{aligned} \tag{26}
$$

$\chi_0(x) = \exp(-T_0(x)V_0(x)\bar{\phi}_l^{-1})$ is the exit rate in state x and $T_0(x)$ are the transition probabilities.

B.3 Elasticities

Let the probability that a single consumer wants to book a listing j by σ_{jt}.

$$\sigma(x) = \frac{\exp(\nu(P^*(x), x))}{1 + \sum_x^X s^*(x)\exp(\nu(P^*(x), x))} \tag{27}$$

The own-price elasticity of listing j at time t is given by the following expression.

$$-100\frac{\exp(-\mu\sigma(x))}{1 - \exp(-\mu\sigma(x))}\mu\sigma(x)(1 - \sigma(x))\alpha P^*(x) \tag{28}$$

The semi-elasticity of demand with respect to a good review of listing j at time t is given by the following expression.

$$\frac{-\exp(-\mu\sigma(K+1, N+1, l)) + \exp(-\mu\sigma(K, N, l))}{1 - \exp(-\mu\sigma(K, N, l))} \tag{29}$$

B.4 Expected operating cost

The expected operating cost is given by the following expression.

$$
\begin{aligned}
\mathbb{E}[\phi_l|\phi_l \leq \delta T(x)V(x)](1 - \chi(x)) =&\bar{\phi}_l - \mathbb{E}[\phi_l|\phi_l > \delta T(x)V(x)]\chi(x) \\
=&\bar{\phi}_l - (\bar{\phi}_l + \delta T(x)V(x))\chi(x) \\
=&(1 - \chi(x))\bar{\phi}_l - \delta T(x)V(x)\chi(x)
\end{aligned}
$$

B.5 Expected entry cost

The expected cost of entry is given by the following expression.

$$
\begin{aligned}
\mathbf{E}_\kappa[\kappa|\kappa \leq \delta\mathbb{E}_l[V((0,0,l))]] =&\frac{\bar{\kappa} - Pr(\kappa > \delta\mathbb{E}_l[V((0,0,l))]\mathbb{E}_\kappa[\kappa|\kappa > \delta\mathbb{E}_l[V((0,0,l))]]}{Pr(\kappa \leq \delta\mathbb{E}_l[V((0,0,l))])} \\
=&\bar{\kappa} - \frac{\exp\left(-\delta\mathbb{E}_l[V((0,0,l))]\bar{\kappa}^{-1}\right)\delta\mathbb{E}_l[V((0,0,l))]}{1 - \exp\left(-\delta\mathbb{E}_l[V((0,0,l))]\bar{\kappa}^{-1}\right)} \\
=&\bar{\kappa}\left(1 + \frac{v_a - \lambda}{\lambda}\ln\left(1 - \frac{\lambda}{v_a}\right)\right)
\end{aligned} \tag{30}
$$