The Cost of the Cold-Start Problem on Airbnb

Florian Dendorfer and Regina Seibel*

August 17, 2025

Abstract

On platforms 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. This paper studies how policies aimed at mitigating the cold-start problem, specifically by lowering prices for new products, affect welfare and how much of the welfare effect is due to a change in the speed of social learning. Based on data from Airbnbs located in Manhattan, we estimate a structural model with short-lived consumers and long-lived hosts who make dynamic entry, exit, and pricing decisions. We then simulate a counterfactual reduction in Airbnb's revenue fee for new listings. While this increases total welfare, we find that it does not accelerate social learning. By contrast, if the cost of the fee reduction is passed on to hosts, faster social learning in itself increases consumer welfare by up to 5% and is the main driver of the total welfare gain.

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, offering consumers an extensive range of options. In this environment, peer-to-peer reviews play a crucial role in helping consumers make informed purchasing decisions based on others' experiences. However, reviews are likely misallocated. 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. Unfortunately, this value is not reflected in consumer choices:

*Dendorfer: Competition Bureau Canada. Seibel: Rotman School of Management, University of Toronto. Email: florian.dendorfer@cb-bc.gc.ca, and regina.seibel@rotman.utoronto.ca. We thank Victor Aguirregabiria, Heski Bar-Isaac, Lanier Benkard, Stefan Bühler, El Hadi Caoui, Gregory Crawford, Florian Ederer, Liran Einav, Gautam Gowrisankaran, Philip Haile, Brett Hollenbeck, Mitsuru Igami, Armin Schmutzler, and conference participants at CRESSE 2023, MaCCI 2023, the Swiss IO Day 2023, DICE 2024, the 17th Digital Economics Conference 2025, the Economics of Platforms Seminar, IIOC 2025 and the Alpine IO Symposium 2025 for helpful comments. The views expressed herein are entirely those of the authors and should not be purported to reflect those of the Competition Bureau Canada.

consumers often prefer reviewed over unreviewed products, which might be of poor quality. This creates the "cold-start" problem (Che and Hörner, 2018; Kremer, Mansour and Perry, 2014), where new products are not discovered sufficiently fast. In other words, social learning is inefficiently slow. Platforms like Airbnb and Amazon appear to understand the importance of initial reviews. They often encourage new sellers to lower prices or offer discounts to buyers as a way to incentivize reviews. For example, Airbnb suggests hosts offer a 20% discount to their first guests, while Amazon has a program where sellers offer products for free in return for reviews.[1]

Whether these policies enhance overall welfare is less clear when taking the endogenous decisions of sellers and market equilibrium effects into account.[2] New sellers may already choose to set low prices to attract reviews and gain a competitive advantage, similar to a learning-by-doing strategy (Cabral and Riordan, 1994, 1997). Additionally, interventions aimed at resolving the cold-start problem might distort prices and seller profits, which in turn influences their decision to enter or exit the market. This can impact product variety and overall welfare for reasons beyond social learning.

In this paper, we estimate the welfare effect of market interventions aimed at mitigating the cold-start problem and identify the role faster social learning plays. Specifically, we develop a structural model of the Airbnb market in Manhattan, New York, where hosts make dynamic pricing, market entry, and exit decisions. Since the quality of a listing is initially unknown, peer-to-peer reviews help reveal it over time. Using our model, we explore counterfactual discounts on new listings, sometimes paired with incentives for guests to leave a review after their stay. Discounts take the form of reductions in Airbnb's revenue fee, and are intended to increase demand and accelerate the arrival of the first review. Even though the overall welfare effect is large, this effect mostly stems from changes unrelated to social learning. We also investigate the consequences of hosts, rather than Airbnb, bearing the cost of the revenue fee reduction and find that such policies are more effective at accelerating social learning. In fact, faster social learning alone may increases consumer welfare by up to 5%. This is partly because such a policy amplifies host's incentives to gain their first review.

We begin by illustrating the cold-start problem within a simple two-period model of social learning with two firms: an incumbent and an entrant, whose quality is initially unknown. The products are vertically, but not horizontally differentiated, so consumers always prefer the option with the higher expected quality. We demonstrate that the cold-start problem occurs if the entrant's expected quality is below, but close to the incumbent's quality. While our illustrative model does not incorporate endogenous pricing, we show in Appendix A.1 that accounting for it would not solve the cold-start problem. We conclude the section with a discussion of the role of endogenous entry and exit, as well as of the outside good. In particular, the relation between the cold-start problem and entry and exit incentives is complex. While increasing the speed of social learning for new listings increases the incentive of hosts to enter and stay in the market,

[1] Airbnb:https://www.airbnb.ca/resources/hosting-homes/a/how-to-set-a-pricing-strategy-15, Amazon: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. Vellodi (2024) demonstrates that allowing for endogenous pricing fully alleviates the cold-start problem in his model with endogenous entry.

all else equal, in a market context, such an increase does not happen in isolation. First, it may be achieved by lower entrant prices, which benefit consumers directly. Additionally, depending on which party has to cover the price difference, any policy will result in a reallocation of rents that will affect the entry and exit incentives of sellers in the market. Since consumers care about the variety of available products, such dynamic responses may have large consequences for consumer welfare.

Next, we present an empirical model of the Airbnb market that incorporates supply dynamics. On the demand side, we employ a discrete choice framework in which guests choose which of the available listings to book based on the expected quality of each listing. Guests update their prior beliefs according to Bayes' rule, using the observed number of reviews and ratings. Importantly, guests do not consider the impact of their potential reviews on future guests when making booking decisions. We assume that if guests choose to review, they do so truthfully. On the supply side, there is a pool of hosts, some of whom are inactive. Active hosts face marginal costs and set rental prices taking into account their effect on the probability of receiving a new review, and therefore future demand. Hosts also make entry and exit decisions: active hosts decide whether to remain active or exit the market to avoid operating costs, while inactive hosts can enter at the beginning of each period, incurring an entry cost. Given the large number of Airbnb hosts, we approximate the symmetric Markov Perfect equilibrium using the oblivious equilibrium concept introduced by Weintraub, Benkard and Van Roy (2008).

We estimate demand and supply separately. For the demand estimation, we use the Generalized Method of Moments. Our estimates highlight the substantial impact of ratings both on guests and hosts. On average, a good review increases a listing's occupancy rate by 6.5%, whereas the estimated average price elasticity of -2.2. The impact of reviews is most pronounced for unreviewed listings, where it is more than four times the average effect. For entrant listings without reviews, we estimate a prior quality expectation of 4.04 stars on a five-star scale. The observed mean rating of 4.69 stars reflects a selection effect: high-quality listings are more likely to persist in the market.

For the supply estimation, we use a nested maximum likelihood approach. In the inner loop, the marginal costs are estimated and the model is solved for a given set of entry and operating costs. The outer loop determines the cost parameters that maximize the likelihood of the observed distribution of listing types and reviews. Our estimates reveal substantial entry costs for hosts, averaging approximately half of their lifetime profit. Based on our estimates of marginal and operating costs, we calculate an average profit margin of 16.7%. The model solution yields rental prices, occupancy rates, average ratings and exit rates that closely align with the observed data. Our findings demonstrate that reviews influence host and consumer behavior in a manner consistent with empirical observations.

In our counterfactual analysis, we examine the effects of offering discounts on new, unreviewed listings to encourage faster learning. These discounts, implemented as reduced revenue fees or a revenue fee waiver for consumers, may be coupled with incentives for review provision: For instance, consumers may be asked to commit to providing a review at the time of booking, with the understanding that this commitment is made in exchange for a reduced price.[3] As previously noted, such policies have welfare implications beyond those directly resulting from a change in the speed of social learning. In the case of lower revenue fees, changes in prices and profits that result from lower revenue fees even absent any change in social learning can be subsumed under the consequences of mitigating double marginalization. To evaluate the contribution of each change separately, we decompose the consumer welfare gain into three parts: The social learning effect, the price effect, and the variety effect. The variety effect refers to the welfare change resulting from a change in the number of listings and distribution of reviews in response to changes in per-period host profits only.

When Airbnb fully waives the revenue fee for new listings, total welfare increases by over 4%, with consumer welfare rising by 6%. While both consumers and hosts benefit, Airbnb experiences a substantial loss in excess of 8% of its revenue, rendering the policy unprofitable for the platform. Moreover, this policy yields unexpected consequences for entrants' pricing decisions, partly counteracting the policy's purpose of accelerating of social learning. This is because the revenue fee waiver puts hosts in an advantageous position, but only as long as their listing does not have any reviews. Once they receive their first review, the revenue fee will be reinstated, and their profit will drop. Increasing their price and slowing down the accumulation of the first review increases the probability that they can reap the benefits of the revenue fee waiver for a longer period. At the same time, the waiver makes new listings with no reviews very profitable, thereby enhancing entry incentives. The increased number of listings, combined with only marginally lower prices, results in lower occupancy rates for entrants and, in effect, slows the learning process. This manifests itself as a negative social learning effect of -0.1% (though small in magnitude). When the revenue fee waiver is coupled with a commitment by guests to provide a review after their stay, total welfare increases by about 5% and consumer welfare increases by more than 6%. Receiving a review with certainty after each booking compensates for the lower expected occupancy rates, resulting in a large social learning effect of 4.6%. Additionally, the policy is profitable for Airbnb.

In our second counterfactual, we explore the case where Airbnb finances the revenue fee reduction through a type-specific lump-sum tax on entrant hosts. Unlike the first set of counterfactuals, this approach enhances hosts' dynamic incentives to acquire the first review, as the lump-sum tax reduces per-period profits of new listings below their status quo level. Consequently, price reductions are two to three times greater than when Airbnb incurs the revenue fee loss, and occupancy rates increase. However, these profit reductions weaken entry incentives, necessitating that lower entry be compensated by lower prices and faster learning to be welfare-increasing. As a result, welfare effects are generally smaller. Consumer welfare rises by 0.9%, nearly half of which is attributable to social learning. When consumers leave reviews after their stay in exchange for a price discount, the gain in consumer welfare increases to more than 4 percent. This counterfactual policy yields a social learning effect of almost 5%, partly due to the alignment of hosts' dynamic pricing incentives with the policy's objective of speeding up social learning.

We conclude that policies that transfer the cost burden to hosts, rather than Airbnb, are more

[3] This is similar to the Amazon Vine Program, where selected consumers can receive free products with the understanding that they will provide a high-quality truthful review.

effective at mitigating the cold-start problem. This finding may explain the prevalence of such policies on digital platforms, exemplified by Amazon's Vine program. Our research demonstrates not only that the cold-start problem occurs on Airbnb, but also that interventions aimed at alleviating it invariably impact market prices and product variety, potentially yielding unintended consequences for consumers. Consequently, platforms must carefully consider and balance these multifaceted effects when designing appropriate policies. This consideration extends beyond Airbnb to any digital platform that relies on peer-to-peer reviews. Our study presents a flexible and tractable model for analyzing various policies targeting the speed of social learning, which can be easily adapted to the wide range of digital platforms we see in practice.

The paper is organized as follows. Section 2 provides a comprehensive review of the relevant literature. In Section 3, we illustrate the cold-start problem in a simplified 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 Section 9, we describe our 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. Kovbasyuk and Spagnolo (2023) and Vellodi (2024) incorporate market entry and find that the cold-start problem acts as an entry barrier. Both papers suggest that limiting information on established sellers may alleviate these barriers and increase surplus.[4] To our knowledge, we are the first to provide an empirical model of social learning through reviews and estimate welfare effects of inducing consumers to explore new products. 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 setup 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 welfare 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

[4] Butler, Carbone, Conzo and Spagnolo (2020) support this intuition experimentally in a public procurement context.

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 underprovided. Airbnb has also been the subject of extensive structural modeling. Most of these studies focus on the housing market (Almagro and Domínguez-Iino, 2025; 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 reviews have 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 Huang (2022), Williams (2022) and Hortaçsu, Natan, Parsley, Schwieg and Williams (2023), dynamic pricing arises from intertemporal price discrimination involving perishable products. In other settings, dynamic pricing reflects the notion that price does not only influence 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 the prices moderate the speed at which consumers learn about products. 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 the oblivious equilibrium concept developed in Weintraub et al. (2008) and Weintraub, Benkard and Van Roy (2010) which is particularly suitable in our setting of many small firms. The oblivous equilibrium has been applied in various studies (Bao et al., 2024; Brancaccio, Kalouptsidi and Papageorgiou, 2020; Chen and Xu, 2023; Frechette, Lizzeri and Salz, 2019).

3 Illustration of the cold-start problem

In this section, we illustrate the cold-start problem in a simple model to provide a foundation for our subsequent analysis and clarify the notions of beliefs and quality in our context.

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 $\omega_j \in [0, 1]$. Quality ω_j reflects the probability with which consumption of j is a "success" and yields utility of one. If it is a "failure", its utility is zero. The quality ω_I of I is publicly known. However, 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) to consumer 2. Note that consumer 1's experience follows a Bernoulli distribution with the probability of a success being equal to ω_E. Conditional on 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.

In this setup, the cold-start problem can only occur if E is expected to be of weakly lower quality than I, i.e. $\frac{a}{a+b} \leq \omega_I$. Then, the consumer will buy the incumbent's product in both periods. However, inducing the consumer in the first period to purchase E may lead the consumer in the second period to learn that E is of higher quality than I. If both consumers choose product I, the consumers welfare is $2\omega_I$. If consumer 1 buys E and consumer 2 buys E, only if consumer 1 reported a successful experience, and I otherwise, the consumer welfare will be $\omega_I + \frac{a}{a+b}\left(1 + \frac{a+1}{a+b+1} - \omega_I\right)$. Hence, changing consumer 1's choice from I to E will lead to an increase in total consumer welfare if and only if:

$$\omega_I \leq \frac{a}{2a+b}\frac{2a+b+2}{a+b+1} := \bar{\omega}_I \tag{2}$$

It is straightforward to see that the inequality holds if $\omega_I = \frac{a}{a+b}$, and by continuity, also when $\omega_I = \frac{a}{a+b} + \epsilon$ for ϵ sufficiently small. However, it does not hold for arbitrarily large ω_I. In particular, $\bar{\omega}_I < \frac{a+1}{a+b+1}$. In other words, the fact that consumer 1 has had success must induce a belief that the quality of E is sufficiently higher than the quality of I.

The illustration presented above abstracts from horizontal differentiation, market dynamics, prices, and the existence of an outside good. In Appendix A.1, we demonstrate that within a standard logit model with horizontal differentiation, endogenous pricing alone is insufficient to resolve the cold-start problem. While a firm can use its price to encourage consumers to discover its product, its incentives do not fully align with those of a social planner. This is because horizontal differentiation in combination with a uniform price does not allow for full rent extraction: Since firms cannot set individual prices, there is a welfare loss even absent Airbnb's revenue fees. Since host profits are not aligned with social welfare, neither are prices.

Dynamic decisions of market participants, such as entry, exit, and pricing, all respond to and influence the rate of social learning. However, incorporating these dynamics complicates the model and, to our knowledge, has not been done in a rich model featuring horizontal differentiation. We fill this gap by developing and estimating an empirical model of social learning, tailored to Airbnb, that accounts for such endogenous, forward-looking host behavior. We evaluate counterfactual price reductions for new, unreviewed Airbnb listings. Such policies, commonly observed in practice, shift the demand from incumbent to entrant listings, thereby increasing the probability of new listings receiving their first review. Absent any other changes, this would result in faster social learning and better informed decisions by consumers. In a model with endogenous entry and exit, faster review accumulation also affects the present discounted value of listings, and consequently, host entry and exit incentives. In particular, all else equal, faster social learning should increase the present discounted value of new listings, thus stimulating market entry and dampening exit incentives for these listings.

However, the shift in demand does not happen in isolation. Hence, it is important to recognize that any policy that reduces the price of new listings induces welfare changes that are unrelated to the cold-start problem:

1. Under imperfect competition, firms set prices above marginal costs. Holding market entry and exit fixed, this leads to an inefficiently high share of consumers opting for the outside good. As a result, lowering the overall price level tends to enhance market efficiency and improve consumer welfare.

2. When market entry and exit are endogenous, price changes affect per-period profits and, by extension, market entry and exit. The direction of profit changes for new listings in response to a policy shifting demand from incumbents to entrants depends on the specifics of the policy. An increase (decrease) in per-period profits raises (lowers) firms' incentives to enter the market. Furthermore, low prices charged by entrants intensify competition, prompting incumbents to exit at higher rates, to the detriment of consumers.

To isolate the contribution of social learning to the welfare changes induced by the policies we study, we decompose the overall change in consumer welfare into three components: the social learning effect, the price effect, and the variety effect, as detailed in Section 9.1. We report the welfare impact of each component separately.

4 Model

4.1 Setup

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 $\mathcal{J}$ of listings, indexed by j. Not all listings are active, so we define the subset of active listings in period t as $\mathcal{A}_t \subset \mathcal{J}$. Each period, there is a new set $\mathcal{I}_t$ of guests, indexed by i. 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 and are risk-neutral. Let u_{ijt} denote a guest i's indirect expected utility of renting Airbnb listing $j \in \mathcal{A}_t$ in time period $t \in \{1, .., +\infty\}$.

$$u_{ijt} = \beta + \gamma \frac{a + K_{jt}}{a + b + N_{jt}} + (1 + f)\alpha p_{jt} + \xi_{jt} + \zeta_{igt} + (1 - \sigma)\epsilon_{ijt} \tag{3}$$

β is the intercept coefficient which only applies to Airbnb listings. γ reflects the utility value of having a successful stay, and α is the rental price coefficient, while p_{jt} is the rental price and f is the fee Airbnb levies on consumers. ξ_{jt} captures unobserved (to the econometrician) listing characteristics. We allow for different levels of substitutability between Airbnb listings among each other, and between Airbnb listings and the outside good by placing them in different nests. ζ_{igt} is the idiosyncratic group preference, and σ is the nesting parameters. ϵ_{ijt} is the idiosyncratic taste shock which is assumed to be distributed i.i.d according to a Type-I extreme value distribution, and ζ_{igt} is distributed such that $\zeta_{igt} + (1 - \sigma)\epsilon_{ijt}$ is also extreme value distributed. 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 assume that the unobserved demand shock is additively separable in a time-constant listing characteristic ξ_j and an i.i.d. time-varying demand shock $\tilde{\xi}_{jt}$, i.e. $\xi_{jt} = \xi_j + \tilde{\xi}_{jt}$. Note that the transitory demand shock $\tilde{\xi}_{jt}$ affects prices and quantities in period t and thereby indirectly the transition probability. However, in any period, the transitory demand shocks from previous periods contain no relevant information about demand conditional on the current state. Hence, the state x_{jt} of listing j in t is $(K_{jt}, N_{jt}, \xi_j) \in X$ where $X = \{(K, N, \xi) \in \mathbb{N}_0^2 \times \mathbb{R} : K \leq N, N \leq \bar{N} \in \mathbb{N}_0\}$. We can write $u_{ijt} = \nu(p_{jt}, x_{jt}, \tilde{\xi}_{jt}) + \zeta_{igt} + (1 - \sigma)\epsilon_{ijt}$ To keep the supply estimation tractable, we will later winsorize the number of reviews at 50, hence $\bar{N} = 50$, and discretize the unobserved listing characteristic parameter ξ_j, see Section 6.2.

We normalize the mean utility of taking the outside good, which we interpret 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.[5] 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 $|\mathcal{I}_t|$ follows a Poisson process with mean μ_t, which may vary over time. Listings are capacity-constrained; at most, one consumer can rent listing j in t. 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 consume the outside good. Hence, demand for Airbnb listing j at time t is given as follows:

$$ccp_j(p_t, x_t, \tilde{\xi}_t) = \frac{\exp\left(\frac{\nu(p_{jt}, x_{jt}, \tilde{\xi}_{jt})}{1-\sigma}\right)}{D_{gt}^{\sigma}(1 + D_{gt}^{1-\sigma})}, \quad \text{where} \quad D_{gt} = \sum_{j' \in \mathcal{A}_t} \exp\left(\frac{\nu(p_{j't}, x_{j't}, \tilde{\xi}_{jt})}{1 - \sigma}\right). \tag{4}$$

$$q_j(p_t, x_t, \tilde{\xi}_t) = 1 - \exp\left(-\mu_t ccp_j(p_t, x_t, \tilde{\xi}_t)\right) \tag{5}$$

[5] 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, see Conlon and Mortimer (2013).

where p_t, x_t and $\tilde{\xi}_t$ are vectors containing all prices, states, and transitory demand shocks of firms active in period t.

State transitions. – If a listing j is booked in t, with probability $v_r(x_{jt}) \in [0,1]$, the guest accurately reports its experience (success, failure) in a review. We refer to $v_r(x_{jt})$ as the conditional review probability (as it is conditional on a booking). [6] The probability ρ_{j0} that a listing's state does not change, captures both the case where the listing is not booked as well as the case where the guest fails to leave a review. The listing receives a good review with probability ρ_{jg} or a bad review with probability ρ_{jb}, depending on the listing's quality prior. Hence $\rho_{jg} + \rho_{jb} = 1 - \rho_{j0}$ is the unconditional review probability.

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 1.

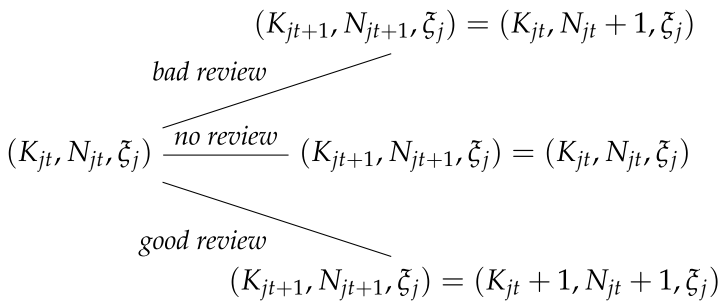

Figure 1: State transitions.

Equation (6) summarizes the transition probabilities.

$$
\begin{aligned}
&\rho_{j0}(p_t, x_t, \tilde{\xi}_t) = 1 - v_r(x_{jt}) q_j(p_t, x_t, \tilde{\xi}_t) \\
\text{and} \quad &\rho_{jg}(p_t, x_t, \tilde{\xi}_t) = v_r(x_{jt}) q_j(p_t, x_t, \tilde{\xi}_t) \frac{a + K_{jt}}{a + b + N_{jt}} \\
\text{and} \quad &\rho_{jb}(p_t, x_t, \tilde{\xi}_t) = v_r(x_{jt}) q_j(p_t, x_t, \tilde{\xi}_t) \left(1 - \frac{a + K_{jt}}{a + b + N_{jt}}\right)
\end{aligned} \tag{6}
$$

The host's problem. – For the remainder of the paper, it is convenient to define the industry s_t as a vector that contains, for each state $x \in X$, the number of firms in state x at time t. Hence, we can write $q_j(p_t, x_t, \tilde{\xi}_t) = q_j(p_t, x_{jt}, s_t, \tilde{\xi}_t)$. The state space is then $\mathcal{S} = s \in \mathbb{R}_0 : \sum_x s(x) < \infty$. Each host may only operate a single listing.[7] Consequently, the full set of listings is the full set of (both active and inactive) hosts $\mathcal{J}$. An active host earns expected variable per-period profit $\pi(p_{jt}, x_{jt}, s_t, \tilde{\xi}_{jt}, c_{jt}) = \mathbb{E}_{\tilde{\xi}_{-j}, p_{-j}} \left[(\text{period length in days}) q_j(p_t, x_{jt}, s_t, \tilde{\xi}_t)(p_{jt} - c_{jt})\right]$, where c_{jt} is the marginal cost of renting out a listing. In addition, a host incurs operating cost ϕ_{jt}, which is a fixed cost. ϕ_{jt} 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_{jt} from entering the market and becoming active as the host must convert their

[6] Note that we assume that reviews are truthful and, while we allow the probability of leaving a review to depend on the current state, it 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] 90% of hosts in our dataset operate a single listing. The average number of listings per host is 1.10.

apartment to an Airbnb rental. All costs are random variables that may vary over time. Each period, c_{jt} is i.i.d. drawn from a distribution G, $c_{jt} \overset{\text{i.i.d.}}{\sim} G$, such that $E[c_{jt}] = \bar{c}_j$. We allow for the fact that entry and exit decisions may happen less frequently than decisions on prices (e.g., because of contractual restrictions such as rental contracts). Specifically, we assume that, in each period, an active listing can make an exit decision and draw an operating cost only with probability r_e, and an inactive listing can make the entry decision and draw an entry cost with that same probability. Conditional on drawing any operating or entry cost, ϕ_{jt} and κ_{jt} follow exponential distributions with mean $\bar{\phi}_j$ and $\bar{\kappa}_j$ respectively. That is, $\kappa_{jt} \overset{\text{i.i.d.}}{\sim} Exponential(\frac{1}{\bar{\kappa}_j})$ and $\phi_{jt} \overset{\text{i.i.d.}}{\sim} Exponential(\frac{1}{\bar{\phi}_j})$. For ease of exposition, we will drop the j subscript from the mean cost parameters in the remainder of this section.

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

1. With probability r_e, inactive hosts observe an i.i.d. entry cost draw κ_{jt} and decide whether to enter at the beginning of $t+1$. They pay κ_{jt} if they enter.

2. Review outcomes are determined and active hosts observe their own state x_{jt}. With probability r_e they observe an i.i.d. cost draw of operating the listing, ϕ_{jt}. Conditional on the draw, they decide whether to stay active or exit.

3. If they stay active, they learn the industry state s_t and pay ϕ_{jt}.

4. Incumbent hosts observe an i.i.d marginal cost draws c_{jt} as well as a transitory demand shock $\tilde{\xi}_{jt}$. They set their price taking into account the implications for future profits. They collect variable per-period profit $\pi(p_{jt}, x_{jt}, s_t, \tilde{\xi}_{jt}, c_{jt})$.

Note that our timing assumption implies that firms make exit decisions knowing their own state realization and that they may exit before the pricing decision.[8] They learn the industry state after all other active firms have also made their exit decision, if any, which is the payoff-relevant state.

4.2 Equilibrium Concept

We will first provide a full description of the strategy space and the value functions which form the baseline of any Markov perfect equilibrium (MPE). We then apply the oblivious equilibrium (OE) introduced by Weintraub et al. (2008) as an approximation of the MPE.

Similar to Weintraub et al. (2008), we assume that the equilibrium is symmetric and features real-valued functions that translate states into actions. We also make use of the fact that the exit strategy follows a cut-off rule: a firm j at time t exits if and only if $\phi_{jt} \leq e(x_{jt}, s_{t-1})$. A minor complication is that our price function has to take into account the marginal cost draw and thus our model allows for price heterogeneity across firms in the same state. The price function implies that each firm j at time t will set a price $\mathcal{P}(x_{jt}, s_t, \tilde{\xi}_{jt}, c_{jt})$. Denote the exit strategy as $e : X \times \mathcal{S} \rightarrow \mathbb{R}_+$ and the price strategy as $\mathcal{P} : X \times \mathcal{S} \times \mathbb{R}^2 \rightarrow \mathbb{R}_+$. Denote that strategy as $(\mathcal{P}, e) \in (\mathcal{M}, \mathcal{E})$, where $\mathcal{M}$ is the set of price strategies and $\mathcal{E}$ is the set of exit strategies. Similarly, denote

[8] We make this assumption to allow for a CCP approach in the marginal cost estimation, see 7.2.

an entry rate function $\lambda \in \Lambda$, where Λ is the set of entry rate functions and $\lambda : X_0 \times \mathcal{S} \to \mathbb{R}_+$ ($X_0 \subset X$ such that X_0 only contains states with $N = 0$). Finally, denote the expected variable per-period profit of firm j, conditional on its competitors following price strategy $\mathcal{P}$ and the firm itself follows strategy $\mathcal{P}'$ as $\mathbb{E}_{c,\tilde{\xi}}\left[\pi(p_{jt}, x_{jt}, s_t, \tilde{\xi}_{jt}, c_{jt}) | p_{jt} = \mathcal{P}', p_{-jt} = \mathcal{P}\right] = \pi_{\mathcal{P}',\mathcal{P}}(x_{jt}, s_t)$. Hence, the value of an active firm just before the realization of the marginal cost demand shock (stage 4 in the timeline above), given that it follows the strategy $(\mathcal{P}', e')$, while competing firms follow the strategy $(\mathcal{P}, e)$ and the entry rate is λ, is as follows:

$$\begin{aligned} V(x_{jt}, s_t|\, \mathcal{P}', \mathcal{P}, e', e, \lambda) = \pi_{\mathcal{P}',\mathcal{P}}(x_{jt}, s_t) + \\ \delta\, \mathbb{E}_{x_{jt+1},s_{t+1}}\left[V(x_{jt+1}, s_{t+1}|\mathcal{P}', \mathcal{P}, e', e, \lambda) - r_e Pr(\phi_{jt+1} < e'(x_{jt+1}, s_t))\mathbb{E}\left[\phi_{jt+1}|\phi_{jt+1} < e'(x_{jt+1}, s_t)\right]\right] \end{aligned} \tag{7}$$

$\delta \in (0, 1)$ 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.

Following Weintraub et al. (2008), we assume that firms make decisions only based on their own state and the long-run industry state. While they make decisions without knowing the current industry state, their decisions collectively give rise to a long-run industry state consistent with their expectation. Hence, with some abuse of notation, we can drop the industry state as argument of the strategy functions $(\mathcal{P}, e)$ and the entry rate λ. If all firms follow a common strategy $(\mathcal{P}, e)$, the state evolves as an independent transient Markov chain. The long-run expected industry state is then $\tilde{s}_{\mathcal{P},e,\lambda}(x) = \lim_{t\to\infty} \mathbb{E}_{\mathcal{P},e,\lambda}[s_t(x)] \forall x$. The corresponding oblivious value function is given as follows:

$$\begin{aligned} V(x_{jt}|\, \mathcal{P}', \mathcal{P}, e', e, \lambda) = \pi_{\mathcal{P}',\mathcal{P}}(x_{jt}, \tilde{s}_{\mathcal{P},e,\lambda}) + \\ \delta\, \mathbb{E}_{x_{jt+1}}\left[V(x_{jt+1}|\mathcal{P}', \mathcal{P}, e', e, \lambda) - r_e Pr(\phi_{jt+1} < e'(x_{jt+1}))\mathbb{E}\left[\phi_{jt+1}|\phi_{jt+1} < e'(x_{jt+1})\right]\right] \end{aligned} \tag{8}$$

For simplicity, we denote $V(x|\mathcal{P}^*, e^*, \lambda^*) = V(x|\mathcal{P}^*, \mathcal{P}^*, e^*, e^*, \lambda^*)$. Definition 1 lists the oblivious equilibrium conditions in our context.

Definition 1. *The oblivious equilibrium is given by the following conditions:*

1. $V(x|\mathcal{P}^*, e^*, \lambda^*) \geq V(x|\mathcal{P}', \mathcal{P}^*, e^*, e^*, \lambda^*) \qquad \forall x \in X, \forall \mathcal{P}' \in \mathcal{M}$
2. $e^*(x) = V(x|\mathcal{P}^*, e^*, \lambda^*) \qquad \forall x \in X$
3. $\lambda_j^* = 1 - \exp\left(-\delta W_j \bar{\kappa}^{-1}\right) \qquad \forall j \in \mathcal{J}$,
   where $W_j = V((0,0,\xi_j)|\mathcal{P}^*, e^*, \lambda^*) - r_e Pr(\phi_{jt} < e^*((0,0,\xi_j)))\mathbb{E}\left[\phi_{jt}|\phi_{jt} < e^*((0,0,\xi_j))\right]$

We refer the reader to Weintraub et al. (2008) for proofs on the existence of the cut-off rule as the optimal exit strategy and the existence of a stationary distribution over states in the environment described above. The oblivious equilibrium is a close approximation of the Markov Perfect Equilibrium in large, unconcentrated markets, where equilibrium industry states exhibit little variation and "observing the industry state and designing strategies [...] do not lead to significant increases in profit" for individual firms (p. 1248). We believe this is an adequate description of the market in our application, with there being hundreds of Airbnb listings in the market at any given time.

In the remainder of this paper, we will focus exclusively on the equilibrium value and strategy functions, and therefore omit $\mathcal{P}^*$, e^* and λ^* for simplicity. Hence, we denote $V(x) = V(x|\mathcal{P}^*, e^*, \lambda^*)$. Note that, in the oblivious equilibrium, the cut-off value for the firm's exit decision is $e^*(x) = V(x)$. Moreover, the exponential distribution of the operating cost allows us to provide a closed form solution of the exit rate $\chi(x_{jt}) = Pr(\phi_{jt} > e^*(x_{jt}))$, where $\chi(x) : X \to \mathbb{R}_+$:

$$\chi(x_{jt}) = \exp(-V(x_{jt})\bar{\phi}^{-1}) \tag{9}$$

It follows from Equation (9) that there is a one-to-one mapping between exit rates and expected values after the operating cost realization but before the marginal cost shock realization. We will exploit this fact when setting up our estimation routine, as we will see in Section 7.2.

Our distributional assumptions allow us to also simplify the firm value at the beginning of each period before the exit decision:

$$V(x_{jt}) - r_e Pr(\phi < e^*(x_{jt}))\mathbb{E}\left[\phi|\phi < e^*(x_{jt})\right] = V(x_{jt}) - r_e(1 - \chi(x_{jt}))\bar{\phi}. \tag{10}$$

Hence, we can rewrite the oblivious value function in the oblivious equilibrium as follows:

$$V(x_{jt}) = \pi_{\mathcal{P}^*,\mathcal{P}^*}(x_{jt}, \tilde{s}_{\mathcal{P}^*,e^*,\lambda^*}(x_{jt})) + \delta\, \mathbb{E}_{x_{jt+1}}\left[V(x_{jt+1}) - r_e(1 - \exp(-V(x_{jt+1})\bar{\phi}^{-1}))\bar{\phi}\right] \tag{11}$$

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.[9] 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 between 2016 and 2019, the time period we analyze, New York City was by far Airbnb's largest market among US cities. Due to the large number of *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 they stayed at 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 (2019), a data analytics company, on all Airbnb listings in Manhattan, New York, between January 2016 and December 2019.[10] 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,

[9] See https://news.airbnb.com/about-us/.
[10] See https://www.airdna.co/.

right). We choose the time frame for two reasons: First, we wanted to exclude the period affected by the COVID-19 pandemic since the tourism industry was one of the most affected industries. Second, on September 5th, 2023, the City of New York mandated that short-term rental must register with the Mayor's Office of Special Enforcement.[11] As a result Airbnbs have all but disappeared as a form of accommodation for tourist.[12] Focusing on a time frame before early 2020 ensures that our analysis is not distorted by these major global or industry events.

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-weekly frequency.[13] 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 suspect that, due to how AirDNA web scrapes the AirBnB website, the occupancy rate around the time a listing first becomes active is underreported. This is because in the month a listing becomes active it is treated as available, even though it might not be active yet. To address this issue, we drop observations that occur prior to the first booking of a listing, ensuring that we only consider data after the listing has definitively become active. Note that this approach excludes 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 \$550, 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 5.7 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.[14]

To estimate the supply-side of the model, we aggregate the data on a weekly level. A week is a sufficiently short time span to ensure a low probability of multiple bookings in one period; the average reservation length of a listing which is around 6 days. Moreover, there are very few weeks with more than one new review coming in for any listing (this only happens in around 5% of listing weeks in our data). Hence, we believe that aggregating at the weekly level is in line with the model assumption of receiving at most one review per period. We chose a higher level of aggregation for the demand estimation to limit measurement error in the market shares we measure.[15] After cleaning the data, we have 161,359 observations at the listing-week level.

Table 1 summarizes key variables in the data on the listing-week level. The average rental

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

[12] 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.

[13] 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.

[14] This leaves us with 7,803 unique listings and 2,263,124 listing-days, excluding blocked ones.

[15] Months refer to four-week periods and do not necessarily correspond to calendar months.

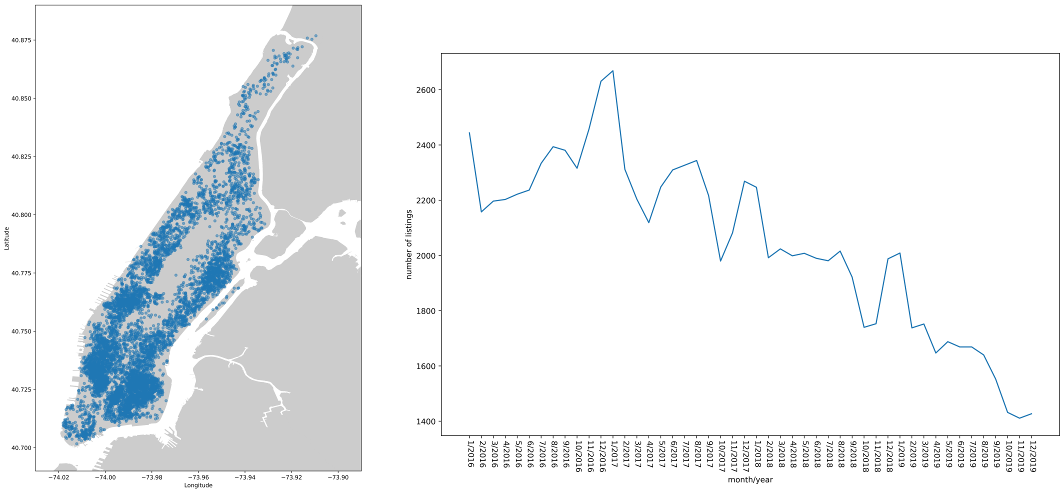

Figure 2: Number of listings over time and their location.

	N	mean	std	min	25%	median	75%	max
Rental rate (in $)	161,359	195.94	62.28	68.63	152.70	187.76	278.02	584.03
Occupancy rate (in %)	161,359	59.36	39.14	0.00	14.29	71.43	100.00	100.00
Number of reviews	161,359	24.99	37.93	0.00	2.00	10.00	70.00	440.00
Rating	139,909	4.69	0.53	1.00	4.60	4.80	5.00	5.00
Weekly share of firms exiting (in %)	208	2.99	3.96	0.00	0.41	1.18	9.03	17.60
Weekly share of new entrants (in %)	208	2.59	3.17	0.00	0.17	1.18	7.89	12.71
Lifespan (in weeks)	6,638	66.49	66.15	1.00	8.00	36.00	156.00	208.00

Table 1: Data summary.

price is $195.94 in our sample. Rental prices exhibit moderate variation. The majority of listings are priced between $150 and $280. On average, Airbnb listings are occupied 59.36% of the time and have been reviewed about 25 times. On average 2.99% of hosts in any given week exit and are no longer active in the subsequent week, while 2.59% have entered.[16] The average lifespan within our sample is 1.3 years, but most listings are active for less than a year.[17]

At 4.69 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 3 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.

[16] 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.

[17] Here, we measure lifespan as the number of weeks for which we observe a listing in our sample. Hence, it has a natural maximum of 4 years, or 208 weeks.

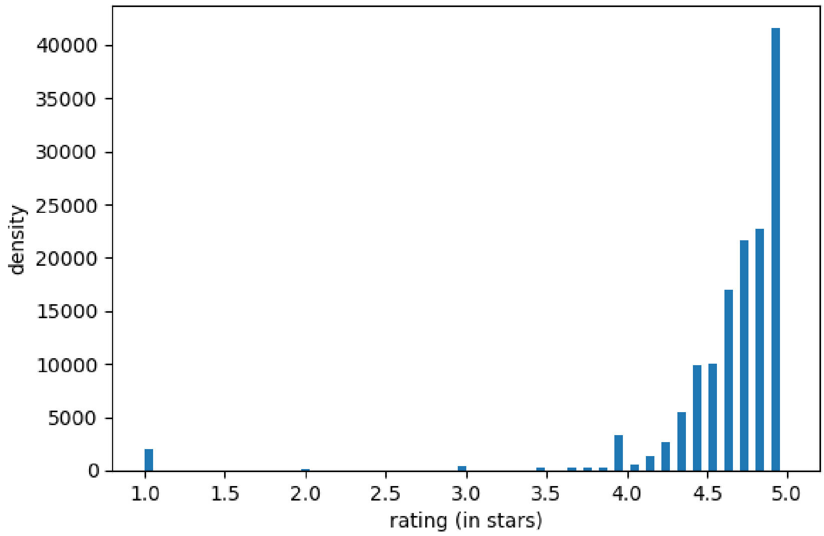

Figure 3: Rating distribution in the data.

To construct the state space, we back out 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 about the number of (good) reviews by filling in the most recent observed value. To keep a reasonably sized state space, we will censor the number of total reviews in the model at 50, i.e., $\bar{N} = 50$. For N exceeding 50, we find that changes in the posterior mean in response to additional reviews become negligible. For example, suppose the prior quality distribution has a mean of 4.04 stars and a variance of 0.27.[18] Figure 4 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 raises its expected quality by 13.5%, whereas the expected quality of a listing with 10 good reviews would only increase by 0.23%, and with 50 good review the posterior mean would increase a mere 0.01%.

We plot the average occupancy rates, rental prices, number of listings, and exit rates depending on N and K in Figure 5. Occupancy rates tend to be larger for listings with many reviews, assuming these reviews are good (Figure 5, top left). Listings without reviews, for example, are more than 30% less likely to be booked than listings with 50 good reviews. The same applies to rental prices, though the relationship is less stark. Listings with 50 good rather than zero reviews have a 11% higher price on average.[19]

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 around high-K states (see Figure 5, bottom left). The right-skewed listing distribution of firms

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

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

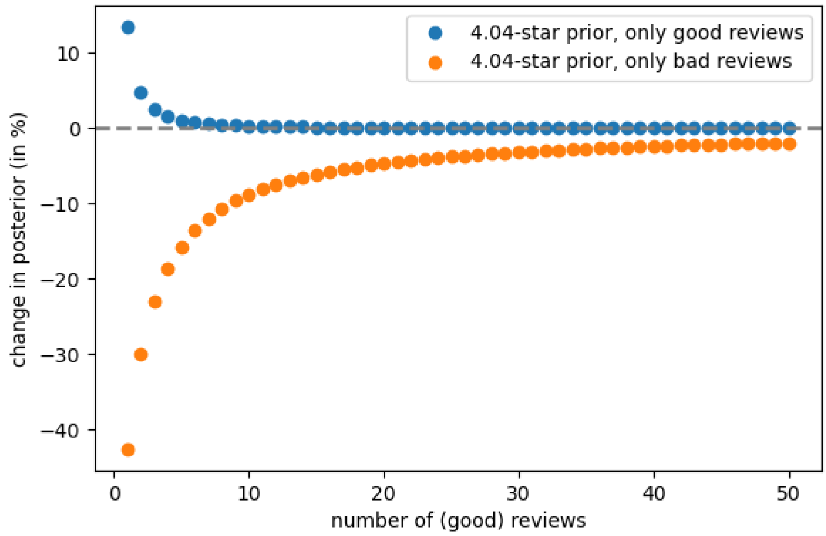

Figure 4: Example posterior changes.

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, around 14% of listings have no reviews and about 18% of reviewed listings have the maximal review count. We also observe the selection effect in hosts' exit behavior (see Figure 5, bottom right). Compared to a reviewed listing, an unreviewed listing is around three times as likely to exit in a given week.

To support the graphical intuition of Figure 5, we regress the rental price, occupancy rate, and exit decision on the number of good reviews and the total number of reviews. Table 2 indicates that, within a listing, a higher number of good reviews is associated with a higher price, while bad reviews do not appear to have a significant impact on price. Similarly, more good reviews correlate with higher occupancy rates, whereas a larger number of bad reviews is associated with lower occupancy. As expected, the exit decision is negatively related to good reviews and positively to bad ones. Naturally, these results only reflect correlations, as all involved variables are equilibrium outcomes.

We combine the main AirDNA dataset with two additional data sources to estimate the market size and the number of potential Airbnb listings in Manhattan. We allow the average daily number of guests arriving in the market to vary by month, based on occupancy data for New York City from NYC & Company.[20] Between 2016 and 2019, New York City sold between 2 and 3.5 million hotel nights per month, with approximately 84% of these bookings occurring in Manhattan. We assume that one in three travelers seeking accommodation consider Airbnb, which roughly aligns with the reported 4% average market share across 50 U.S. cities in Farronato and Fradkin (2022). The total number of potential listings (J) is based on the 2017 New York City

[20] See https://assets.simpleviewinc.com/simpleview/image/upload/v1/clients/newyorkcity/FYI_Hotel_reports_March_2020_9afc1fca-79c1-47a9-8a21-fdefa3c4ffcb.pdf.

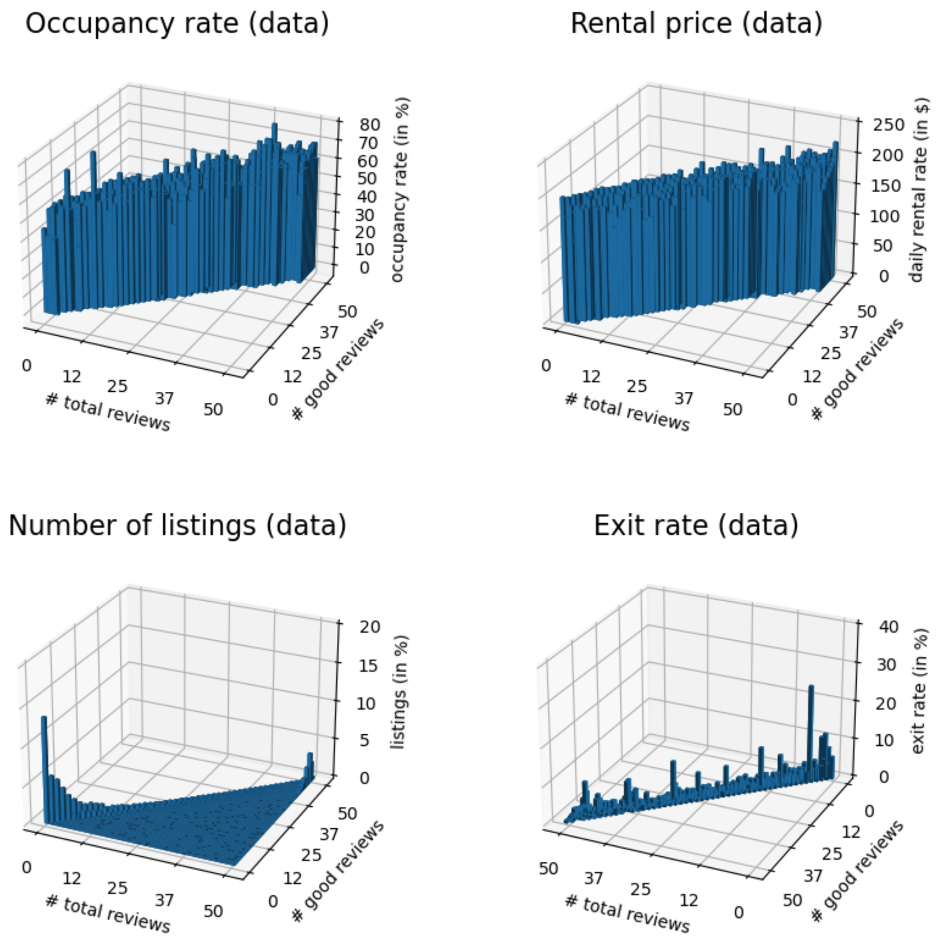

Figure 5: Occupancy rate (top left), rental rate (top right), number of listings (bottom left), and exit rate (bottom right) by number of (good) reviews in the data.

Housing and Vacancy Survey.[21] By adding the survey estimate of relevant vacant units with the average number of available Airbnbs in the data, we arrive at an estimate of 24,522 apartments that could potentially be converted into Airbnb listings.[22]

6 Demand Estimation

6.1 Estimation Approach

We estimate the parameters governing demand by the Generalized Method of Moments (GMM) on data aggregated to the monthly level. Some listings 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. We invert demand and back out ξ_{jt} to compute the moment conditions:

[21] See https://web.archive.org/web/20241204060848/https://www.census.gov/data/datasets/2017/demo/nychvs/microdata.html.

[22] Specifically, we consider vacant one-bedroom units with fewer than four rooms available for rent in Manhattan.

	Price	Occupancy	Exit
K	0.11***	0.002***	-0.0006***
	(0.03)	(0.000)	(0.0000)
N-K	-0.42	-0.009***	0.0010***
	(0.35)	(0.003)	(0.0003)
Listing FE	Yes	Yes	No
Year × Month FE	Yes	Yes	Yes
Observations	90,032	90,032	90,032

*Robust standard errors in parenthesis: ***: 0.01, **: 0.05, *: 0.1*

Table 2: Rental rate, occupancy rate, exit rate and reviews in the data.

$$\nu(p_{jt}, x_{jt}, \tilde{\xi}_{jt}) = \ln\left(\frac{-\ln(1-B_{jt})}{\mu_t}\right) - \ln\left(1 + \frac{\sum_{j\in\mathcal{A}_t}\ln(1-B_{jt})}{\mu_t}\right). \tag{12}$$

We account for seasonality by allowing the daily arrival rate to vary by month as described in Section 5.

Without instrumental variables, our estimates are biased because the rental price and the number of (good) reviews are endogenous. To address price endogeneity, we use a cost shifter as an instrument for the rental price, and past transitory demand shifters as instrument for the number of reviews. Specifically, we use the average reservation length of a listing to capture the marginal cost of renting it out. The rationale is that longer reservations reduce the host's workload per night, as checking guests in and out and cleaning typically occur only once at the beginning or end of a stay. The instrument would be invalid if reservation length were systematically related to a listing's quality – for instance, if lower-quality listings tended to have longer reservations on average. We argue that this is unlikely. Presumably, guests adjust their demand on the extensive margin (e.g., not picking a low-quality listing) rather than on the intensive margin (e.g., deciding to stay three instead of four days because a listing is of low quality).

For the (good) number of reviews, we use the de-meaned (at the listing level) five-, six- and seven-month lagged occupancy rates, and their interaction with the rating as instruments for the (good) number of reviews. Since we adjust past occupancy rates by the listing's average, these rates reflect transitory demand shocks rather than persistent listing quality. For example, neighborhood festival, holidays, or university graduation ceremonies may temporarily boost occupancy rates in specific locations. The lagged occupancy rate would not be a valid instrument if such demand shocks persisted into the current period. This could occur if a listing or its surroundings experienced extended or permanent changes, observable to guests at the time of booking—for instance, if the apartment was remodeled or the surrounding area was upgraded. However, we believe this issue is minimal in our data, as such changes would likely result in sustained and significant price adjustments. In our sample, 18.8% of listings show no rental price changes, and the standard deviation of rental prices within a listing is only \$5.83.

As is standard, we use the common instrument of the total number of listings active in a given period as an instrument for the within-group market share.

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. Naturally, the coefficient of the rental price is identified by variation in the rental price. 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 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.

6.2 Results

We present our demand estimates in Table 3. Column (1) shows our results for a plain logit model without nests, while, in column (2), we place Airbnb listings in one nest, with the outside good in a different nest. We find that the nesting parameter in column (2) is small and not significant, suggesting that the substitutability between Airbnbs is similar to the one between Airbnbs and the outside good. Nonetheless, we will focus on the point estimates of the nested logit specification in the remainder of this paper and use them in the supply side estimation. Based on the nested logit estimation results, ψ and ι imply a prior mean and variance of 4.04 stars and 0.27, respectively, and the coefficient of the posterior γ is large and significant at the 10% level.[23] Figure 6 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 listing selection; highly rated listings are observed in the data, whereas poorly rated listings are not.

The estimated rental price coefficient corresponds to an average listing-month price elasticity of -2.2. 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 rea-

[23] See Table B.1 in Appendix B.4 for the first-stage regression results of our demand estimation.

		(1) No nest		(2) Airbnb nest	
Constant	β	-10.571	(3.174)	-9.561	(2.136)
Prior	ψ	-1.419	(0.732)	1.152	(1.177)
	ι	1.196	(1.021)	0.296	(1.362)
Rental price	α	-0.023	(0.007)	-0.016	(0.006)
Expected quality	γ	6.068	(2.773)	3.416	(1.945)
Nesting parameter	σ			0.046	(0.110)
Observations		24,816		24,816	

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

Table 3: Demand estimates.

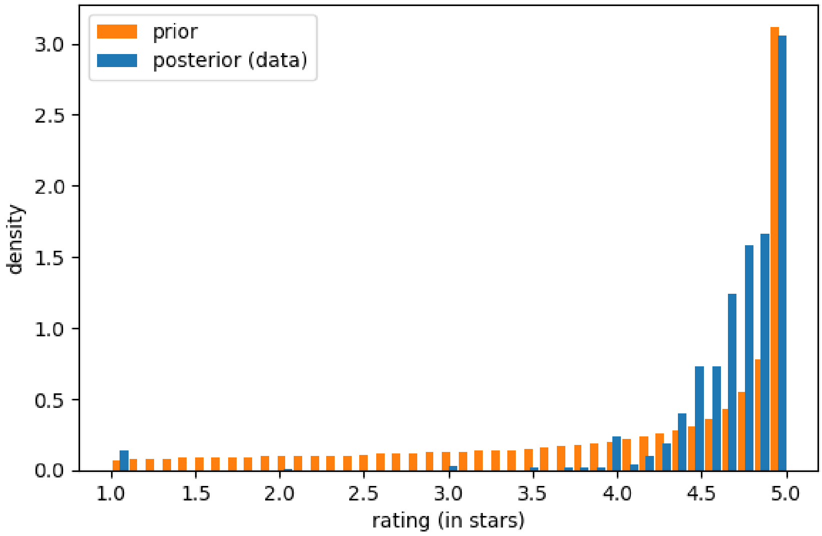

Figure 6: Estimated prior distribution and rating distribution in the data.

sonable. It is possible that our estimate of the price coefficient reflects search frictions; if guests are not aware of all available listings, they would substitute to other options at a comparatively lower rate in response to a price increase. The constant β is estimated to be negative and large in magnitude, reflecting that most guests take the outside good (i.e., book a hotel) rather than book an Airbnb. This is in line with Airbnb's small market share of 4%, as reported in Farronato and Fradkin (2022).

We also compute the semi-elasticity of demand if a listing receives a good review. On average, the occupancy rate increases by 6.46% in response to a good review. Figure 7 shows that the demand for listings with fewer reviews is more elastic. Guests who see an additional review will only slightly adjust their expectations of a listing's quality 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 informa-

	mean	std	min	25%	50%	75%	max
Own-price elasticity	-2.2056	1.3533	-10.4414	-2.8731	-1.9713	-1.2520	-0.0000
Good review semi-elasticity	6.4558	11.9729	0.0000	0.3364	1.1018	5.2236	82.1371

Table 4: Summary of elasticities.

tion available to assess its quality. Accordingly, the average good review semi-elasticity in the data is 28.5 for unreviewed listings but a mere 0.3 for listings with 49 reviews. Table 4 provides summary statistics for the two elasticities we estimate.

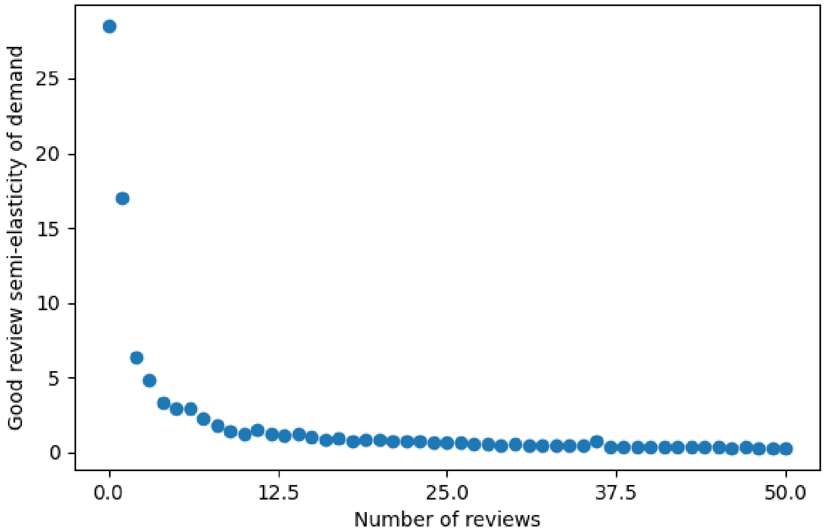

Figure 7: Mean good review semi-elasticity conditional on the number of reviews.

We are not aware of any studies that provide review elasticities for Airbnb to which we can compare our estimates. Fradkin and Holtz (2022) find no effect of reviews on Airbnb reservations, but this result is likely driven by the specifics of their field experiment. For comparison, Reimers and Waldfogel (2021) estimate that book sales increase by approximately 2.6% if the book is reviewed in the *New York Times*.

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* $22.33 higher per day for a listing with one review and a 5-star rating compared to an unreviewed listing. An additional good review of a listing with 49 reviews and a four-star rating is worth only $1.06 per day to guests.

Using our demand estimates, we back out the demand shock ξ_{jt} for each observation. We compute $\bar{\xi}_j$, the average demand shock of each listing. To ensure a manageable state space, we define four listing types instead of allowing state transitions to be listing-specific. We compute the quartiles of the empirical distribution of the estimated $\bar{\xi}_j$s and assign type 1 through 4 to a listing depending on the quartile $\bar{\xi}_j$ falls into. Note that by distinguishing only four types, we choose to be parsimonious and limit the size of the state space in order to keep the model tractable. We report summary statistics for the different types in Table 5.

On average, higher types feature a higher occupancy rate. While a type 4 listing is, on average, booked 64% of the time, the average occupancy rate of a type 1 listing is 49%. The correlation between the type and the rating is very weak (Pearson correlation coefficient: -0.002). The number of reviews tend to increase in the type (0.09). At the same time, the rental price is highly correlated with the listing type (0.47). This correlation provides support for the notion that types are largely differentiated by characteristics that are observable to consumers, rather than the unobserved quality.

	$\bar{\xi}$	Price	Occupancy	Reviews	Rating
type 1	-1.27	$146.25	48.79%	17.00	4.69 stars
type 2	-0.35	$173.74	61.33%	26.83	4.68 stars
type 3	0.24	$202.64	64.06%	29.27	4.70 stars
type 4	1.30	$261.14	63.24%	26.86	4.67 stars

Table 5: Summary: listing types.

7 Supply Estimation

7.1 Calibrated Supply-Side Parameters

We do not estimate all the parameters of our model, but we calibrate some of them using data moments. Recall that a period is one week. For the discount factor δ, we choose 0.9988, implying an annualized interest rate of 6.7%. We allow the conditional review probability v_r to vary by the number of current reviews. Hence we estimate v_r as the inverse of the number of booked weeks required until an additional review is observed in the data for listings with less than 50 reviews. In this way, we determine that v_r varies between 38% and 62% depending on the number of reviews a listing has. It is important to note that we do not allow v_r to vary based on the rating, the rental price, or whether the guest experiences a success or a failure.[24]

Given that rental contracts in the market typically operate on a monthly cycle, it is reasonable to assume that hosts make their entry and exit decisions on a monthly basis. Consequently, we set the probability of making an entry or exit decision decisions in any given week to be 25%, i.e., $r_e = 0.25$.

By the construction of listing types, the distribution of types across active listings is uniform. However, since we do not observe inactive listings, we cannot infer the unconditional distribution of listings. We assume that higher-quality types are less likely to occur; specifically, types three, two and one are twice, thrice, and four times as likely to occur than type four. The average number of listings of a certain type outside the market is J_l minus the total number of listings of that type in the market, where $l \in \{1, 2, 3, 4\}$. While this assumption affects the estimated mean cost parameters, we find that the average *incurred* entry and operating costs

[24] See Figure B.1 in Appendix B.4 for the empirical review probabilities in the data.

we estimate are fairly robust to the assumption of the listings-type distribution among potential listings. Specifically, if the share of a type among the total number of listings is larger (smaller), our average entry cost estimate is larger (smaller) as to offset the larger (smaller) number of potential entrants. Nevertheless, we caution the reader to interpret our estimation results with the assumption we made regarding the listing-type distribution in mind.

We find that the majority of the variance of the overall demand shock ξ_{jt} is explained by the time-constant type $\bar{\xi}_j$. Moreover, the variance of $\tilde{\xi}_{jt}$ is dwarfed by the marginal cost variation based on the estimation routine we describe below. Hence, for computational reasons, we only integrate over the marginal cost shocks in the model solution below and assume $\tilde{\xi}_{jt} = 0$ for all j and t.

Lastly, Airbnb charges – both in the model and in reality – a commission fee of 14.2%.[25]

7.2 Estimation Approach

We estimate the type-specific cost parameters of the model, denoted by $k = (\bar{\kappa}, \bar{\phi}, \bar{c})$, with $\bar{\kappa} = (\bar{\kappa}_1, \bar{\kappa}_2, \bar{\kappa}_3, \bar{\kappa}_4)$, $\bar{\phi} = (\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3, \bar{\phi}_4)$ and $\bar{c} = (\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4)$, conditional on the demand parameter estimates $\theta = \{\beta, \gamma, \psi, \iota, \alpha, \bar{\xi}_{,1}, \bar{\xi}_2, \bar{\xi}_3, \bar{\xi}_4\}$, with a nested Maximum Likelihood routine. In the inner nest, we estimate the marginal costs parameters $\bar{c}$ for a guess of the entry costs $\bar{\kappa}$ and operating costs $\bar{\phi}$ using GMM, and solve the model to arrive at the state distribution implied by the cost parameters. In the outer nest, we search for the entry and operating cost parameters which maximize the likelihood over the state distribution $s^*(x|k, \theta)$.

7.2.1 Inner nest

Marginal cost estimation We will back out marginal cost from the optimality condition for prices.[26] Assuming that competing firms follow the optimal pricing and exit strategies, and the entry rate is individually optimal as well, with some abuse of notation, we denote demand for listing j with $q(p_j, x_j)$. We drop the j subscripts and use Equation (10) to express the firm value at the beginning of each period. Hence, the dynamic first-order condition of firm j can be rearranged as follows:

$$c = p - m(p,x) + \left(\frac{\partial q(p,x)}{\partial p}\right)^{-1} \delta \sum_{x'} \frac{\partial Pr(x'|p,x)}{\partial p} (V(x') - r_e(1-\chi(x'))\bar{\phi}) \tag{14}$$

where $m(p,x)$ is listing j's mark-up, i.e., $-q(p,x)\left(\frac{\partial q(p,x)}{\partial p}\right)^{-1}$. Note that all the components of the above cost equation are jointly identified by the demand parameters and the data, with the exception of $V(x')$. Nonetheless, we are able to estimate the value function conditional on a guess of the operating cost parameter, as we describe below.

We use the fact that, conditional on a candidate of the average operating costs, there is a one-to-one mapping between the exit rate and the expected value of a listing after the operating

[25] See https://www.airbnb.ca/resources/hosting-homes/a/how-much-does-airbnb-charge-hosts-288.
[26] While this is standard for static models, we are not aware of previous work implementing the estimation approach we use for dynamic models of pricing.

cost realization but before the realization of the marginal cost shock and the demand shocks: $\chi(x) = \exp(-V(x)\bar{\phi}^{-1})$. Note that, for states that we observe in the data, we can estimate exit rates non-parametrically. We can then use estimated exit rates $\hat{\chi}(x)$ to represent the value function before the operating cost realization as follows:

$$\hat{V}(x|\bar{\phi}) = -\ln \hat{\chi}(x)\bar{\phi} \tag{15}$$

Hence, the value of a listing in state x at the beginning of a period is identified by the empirical exit rate in that state up to the operating cost parameter. We can replace the value function in Equation (14) by its empirical equivalent:

$$c = p - \hat{m}(p, x) + \sum_{x'} \frac{\partial \hat{P}r(x'|p, x)}{\partial p}(\hat{V}(x'|\bar{\phi}) - r_e(1 - \hat{\chi}(x'))\bar{\phi}) \tag{16}$$

To be consistent with our state space, we parameterize marginal costs as follows: $c_{jt} = c_l \mathbb{1}(type_j = l) + \eta_{jt}$. Based on this specification, our moment conditions are $\mathbb{E}[Z_{jt}\eta_{jt}(\bar{\phi}_{l(j)})] = 0$. Hence we get marginal cost estimates as a function of the operating cost parameter guess: $\hat{c}(\bar{\phi})$.

Solving the Model Given a candidate for the operating cost and entry cost, the above step gives us the marginal costs. Next, for each candidate of the cost parameters $k = (\bar{\kappa}, \bar{\phi}, \bar{c}(\bar{\phi}))$, we solve the model in four steps. We start by formulating the initial guess of the pricing policy function $P_0(x, c)$, the number of listings $s_0(x)$ per state, and the host's value function $V_0(x)$.[27]

Step 1 – Based on the guess, we determine a host's best response $P_1(x, c)$ if all remaining hosts adhere to $P_0(x, c)$. We find $P_1(x, c)$ by combining numerical integration with the method of gradient ascent. We draw marginal costs based on Halton draws from the marginal cost distribution we estimated in the data. We then iterate over Equation (17), where k is the iteration step and $P_1^0 = P_0(x, c)$, until the change in the rental price, $P_1^{k+1} - P_1^k$ for any state x is less than \$0.1.

$$P_1^{k+1}(x, c) = P_1^k(x, c) + w'(x, c|P_1^k, P_0) \tag{17}$$

where w' is the first-order derivative of the host's value after learning the cost shock given the host follows price rule P_1^k while competitors follow price rule P_0. Since hosts don't observe their rivals' marginal cost shocks, they have to form expectations about the rivals' prices based on the marginal cost distribution and the pricing function P_0.[28]

Step 2 – Assuming that all hosts set their prices according to $P_1(x, c)$, we compute the expected variable per-period profit $\pi_1(x)$ and occupancy rate $q_1(x)$ *before the cost shock realization*. The latter forms the baseline for our expected transition matrix $T_1(x)$. We use the expected profit, demand and transition matrix together with the current candidate for the operating cost to compute the value function $V_1(x)$ from Equation (11).

[27] We initialize the model with the empirical state distribution, value functions of 0 for all states, and stationary optimal prices.
[28] The exact expressions can be found in the Appendix B.1.

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_{l1}. Together with the occupancy rates $\mathbf{q}_1$, χ_1 and $\boldsymbol{\lambda}_1$ allow us to compute an expanded transition matrix $\mathbf{F}_1$. We use $\mathbf{F}_1$ to solve for the new, stationary listing distribution $\mathbf{s}_1$, where $\mathbf{s}_1$ is given by:

$$\mathbf{s}_1 = \mathbf{s}_1 \mathbf{F}_1 \tag{18}$$

Step 4 – If the absolute difference between $V_1(x)$ and $V_0(x)$ or $\mathbb{E}_c[P_1(x,c)]$ and $\mathbb{E}_c[P_0(x,c)]$ or $s_1(x)$ and $s_0(x)$ for any x exceeds 0.000001, we update the guess to $P_1(x,c), s_1(x)$, 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).

7.2.2 Outer Maximum Likelihood estimation

Denote the set of states associated with type l by X_l. Equipped with the solution to the model from the previous step, the likelihood function is given by:

$$\begin{aligned} L(k|x_{jt},\theta) = &\sum_t \Bigg\{ \sum_j \left[\sum_{x\in X} \mathbb{1}(x_{jt}=x) \ln(s^*(x|k,\theta)) \right] \\ &+ \sum_{l\in\{1,2,3,4\}} \left(J_l - \sum_{x\in X_l} \sum_j \mathbb{1}(x_{jt}=x) \right) \ln\left(\frac{J_l}{J} - \sum_{x\in X_l} s^*(x|k,\theta) \right) \Bigg\} \end{aligned} \tag{19}$$

The maximum likelihood estimation requires us to solve the model and determine $\mathbf{s}^*$ repeatedly for different cost parameter candidates.

7.3 Results

Type	1	2	3	4
Marginal Cost (in $)	73.73*** (0.87)	66.32*** (1.25)	76.49*** (1.26)	120.37*** (1.25)
Mean Operating Cost (in $1,000)	1.76*** (0.03)	3.01*** (0.07)	3.91*** (0.05)	4.80*** (0.05)
Mean Entry Cost (in $1,000)	260.43*** (0.15)	890.41*** (1.50)	886.80*** (0.71)	543.99*** (1.36)

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

Table 6: Supply estimates.

The estimation results are shown in Table 6. All estimated parameters are highly significant. Note that we estimate the mean parameters of the entry and operating cost distribution, which is by many order of magnitudes larger than entry and operating costs that are actually *paid* upon entry as firms will only enter and incur the cost if the cost draw is low enough. Our estimates imply that hosts on average incur entry costs ranging from $638 to $3,384, which are monotone increasing in type and around $1,713 on average. This compares to an average expected lifetime profit of entrants between $1,278 and $6,790, with a total average of $4,079. Listings of any two types are by construction equally likely observed in the market although higher type listings are more profitable. Our estimates suggest that entry costs are consistently around 50% below the listing value upon market entry.

We estimate that on average hosts pay between $1,222 and $3,779 in operating costs every four weeks, depending on the type, with an average of $2,601. As with entry costs, higher-type listings incur higher operating costs. Intuitively, higher-type listings are more desirable to guests but also the hosts. 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 around three times as much as a type-1 listing to maintain. Nonetheless, hosts earn more profit from higher-type listings. The weekly present discounted values of an active type-1 is with $2,800 more twice as high as of of an active type-4 listing.

While the average marginal costs are also broadly increasing in type, the increase is non-monotonic with type-2 listings having lower marginal costs on average than type 1 listings. Despite this non-monotonicity in marginal costs, total per-period profits, i.e. variable per-period profits less operating costs, are monotonically increasing in type. The average total profit ranges from $248 to $814 with an average of $524. With average revenue per month ranging from $1,824 to $4,150, and a total average of $3,020, this implies an average profit margin of around 16.7% for active listings. As, expected type-4 listings have the highest profit margin with 19.6%, while type-1 listings have the lowest with 13.6%.

8 Model Fit

To assess the accuracy of our model, we compare key empirical moments derived from the data with those arising in the model solution, when parameterizing it with the estimates obtained in the previous section. Table 7 presents a detailed comparison of these moments.

While the model's average number of listings and average prices are somewhat higher than those observed in the data, differences are small. The average occupancy rate in the simulated data aligns well with its empirical counterpart, underscoring the model's capacity to closely match data moments. Similarly, the average rating differs only in the second decimal place. The average number of reviews is somewhat lower in the model (notably after winsorizing data reviews at 50), while the exit rate is slightly higher. We believe that this comparison demonstrates the model's strong ability to capture essential features of the data.

	# listings	Price	Occupancy	# reviews	Rating	Exit rate
Model	784	$207.87	59.32%	11.10	4.67	4.21%
Data	765	$195.94	59.36%	17.64	4.69	2.99%

Table 7: Data versus model averages.

Figure 8 compares the rating distribution in the model solution ("posterior (model)") to the rating distribution in the actual data ("posterior (data)"). The frequency of one-star, three-star, and four-star listings (i.e., ratings comprised of three good reviews and one bad review) are remarkably similar in the model and the actual data. The model predicts more five-star ratings than we see 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 50. Hence, the model allows for little variation in the ratings close to 5 stars. Ratings in the simulated data are based on at most 50 reviews but possibly on many more in the actual data. Note that Figure 8 illustrates the selection effect that underlies the model as well as the actual data. Recall that the mean quality of *entrant* listings is estimated to be 4.04 stars. This is lower than the average rating of *all active* listings, both in the model solution and actual data.

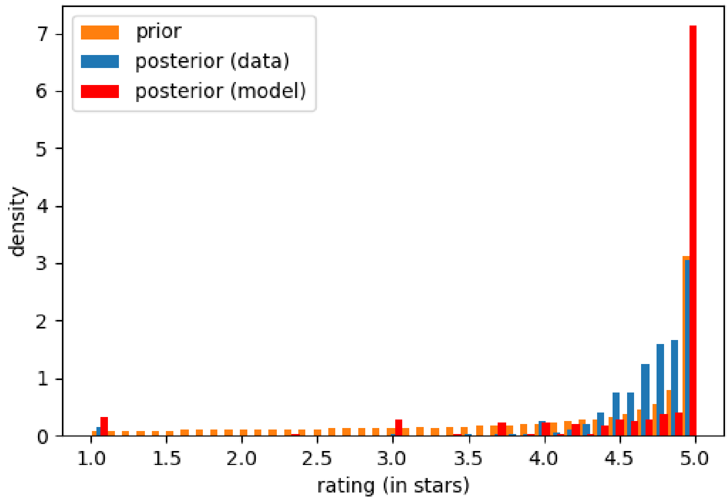

Figure 8: Estimated prior distribution, rating distribution in the data, and model rating distribution.

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 784 active listings, compared to 765 in the data. Figure 9 (bottom left) shows that the number of listings generated by the model for each number of (good) reviews has a similar "pitched-fork" shaped distribution as the observed data, even though the pitchfork in the model might be somewhat more pronounced. Around 17% of active listings are unreviewed in the model, which compares to around 13% in the data. 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. Lastly, Figure 9 shows the model produces rental prices and occupancy rates that are comparable to the actual data (top right and top left, respectively).

Our model effectively captures many key aspects of the data. In particular, the model produces a state distribution similar to the empirical one. Furthermore, the direction and magnitude of the responses of hosts and guests to reviews in the model are broadly consistent with the data. Our counterfactual policies aim to increase the rate at which the quality of new listings is

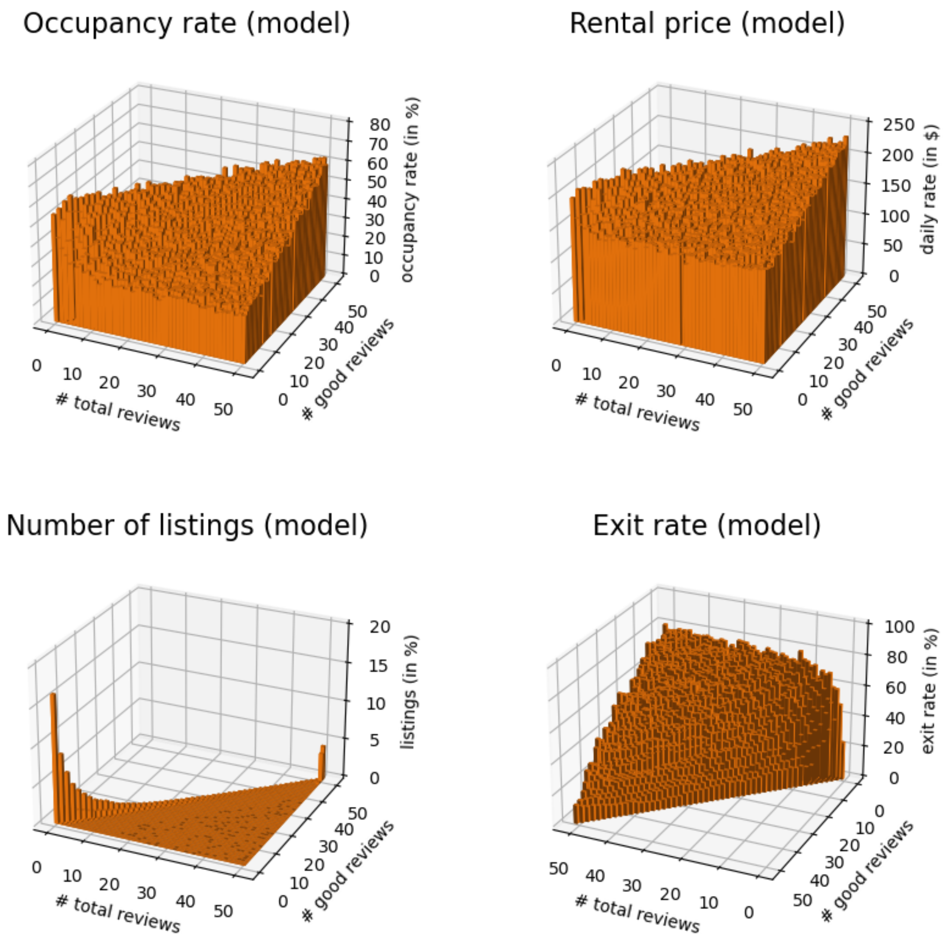

Figure 9: Estimated occupancy rate (top left), rental rate (top right), number of listings (bottom left), and exit rate (bottom right) by number of (good) reviews.

revealed by boosting their occupancy rates. According to our model, new listings underperform older ones economically. The occupancy rate of a new listing is typically 6.2 percentage points below the average, and its rental price is $20 lower than the average. Taken together, new listings generate 8% less revenue than the average listing. Consequently, our model predicts that new listings are around twice as likely to exit the market compared to the average listing. We find a similar relationship in the data: new listings are twice as likely to exit in any given week. Moreover, new listings are booked 24% less frequently and generate 25% less revenue than the average listing.

9 Counterfactual analysis

Having estimated the model parameters, we simulate counterfactual price reduction for new, unreviewed listings. Specifically, Airbnb reduces or fully waives the revenue fee it has historically charged. We consider a version of the policy in which the fee reduction is coupled with incentives for review provision. In addition, we vary who bears the cost of the policy. We consider two cases:

1. Airbnb absorbs the loss from the revenue fee reduction.
2. Entrant hosts pay a type-specific lump-sum tax equal to the expected forgone Airbnb revenue fee.

We calculate the compensating variation of each policy following Small and Rosen (1981) and McFadden (2012). We compute the welfare change that results from the policy as the sum of the change in total host profits (i.e., variable per-period profits less operating costs and entry costs), the change in the Airbnb fee revenue, and the compensating variation, which we refer to as the consumer welfare change in what follows.

9.1 Welfare decomposition

Lowering the prices of new listings have welfare effects beyond those associated with the cold-start problem as briefly discussed in Section 3. In our counterfactual analysis we focus on the specific case of lowering prices by means of reducing the revenue fees on new, unreviewed listings. Since the fee increases the effective marginal cost of Airbnb hosts, reducing or fully waiving it is akin to mitigating the welfare losses from double marginalization: consumers experience lower prices and shifts in both the number and composition of listings, as we explain below. In our decomposition, we separately identify the welfare effects of these changes in addition to those from changing the speed of social learning.

In the following, we focus on consumer welfare when comparing the sizes of these effects, since the cold-start problem is a collective action issue among consumers. Consumer welfare, $CS(s, p)$, is affected by two equilibrium objects, the state distribution s and the equilibrium prices p. We denote the status quo consumer welfare by $CS(s^{SQ}, p^{SQ})$, and the counterfactual consumer welfare by $CS(s^{CF}, p^{CF})$. A reduction in double marginalization immediately affects consumers in two ways that are unrelated to social learning. First, holding entry and exit fixed, lower fees are passed on to guests through lower prices. Second, allowing for endogenous entry and exit, they increase entrant variable per-period profits and strengthen hosts' incentives to enter (or reduce incentives to exit), thereby affecting listing variety. However, this depends on Airbnb absorbing the fee reduction as a loss. If instead Airbnb charges entrant hosts a lump-sum tax to offset the cost, entrant profits and product variety may decline.

Therefore, to isolate the contribution of faster social learning to consumer welfare gains, we decompose the overall effect into three components. First, we identify the state distribution s^{FP}, which would have resulted had variable per-period profits not changed, but only the transition probability. To this end, we solve for the equilibrium state distribution of the model with variable per-period profits being fixed at their status quo levels but the transition probabilities that arose in the counterfactual through fixed point iteration. Second, we evaluate consumer welfare at three different combinations of state distribution and prices to evaluate the three components of the consumer welfare effect $CS(s^{CF}, p^{CF}) - CS(s^{SQ}, p^{SQ})$:

1. The **variety effect** $CS(s^{CF}, p^{CF}) - CS(s^{FP}, p^{CF})$ measures the consumer welfare change due to the variety changes from variable per-period profit changes only.

2. The **price effect** $CS(s^{FP}, p^{CF}) - CS(s^{FP}, p^{SQ})$ captures the consumer welfare change from price changes only, holding variety fixed.

3. The **social learning effect** $CS(s^{FP}, p^{SQ}) - CS(s^{SQ}, p^{SQ})$ isolates the change in consumer welfare change resulting solely from the transition probability changes, holding prices constant.

Below, we report these effects relative to the status quo consumer welfare, $CS(s^{SQ}, p^{SQ})$, to provide a clear sense of their magnitudes.

9.2 Reducing the revenue fee for entrants

9.2.1 Airbnb pays for the revenue fee reduction

In this section, we start by analyzing the welfare implications if Airbnb were to fully waive the revenue fees on new listings. We consider two cases: one where no other changes are made and another where the fee waiver is contingent on the consumer providing a truthful review post-purchase.

Our findings are summarized in Table 8. We find that waiving the revenue fee for unreviewed listings (column (1)) increases total welfare by 4.3%. The welfare gain stems from increases in consumer welfare and total host profits which outweigh the loss in Airbnb's revenue. Interestingly, entrant prices only decrease by 0.8%, despite the full 14.2% revenue fee being waived. This suggests that the reduced fees are largely absorbed by entrant hosts rather than passed on to consumers. Counterintuitively, the occupancy rate of new listings decreases, which is at odds with the objective of incentivizing consumers to explore more. At the same time, the total number of listings increases substantially, by 6.8%.

The surprising occupancy decline warrants further discussion. First, it is important to understand why entrant prices almost fully offset the revenue fee reduction. Without dynamic considerations, the only explanation would have been that hosts have substantial market power. With an estimated own-price elasticity of -2.2, this is only part of the story. Additionally, changes in dynamic pricing incentives seem to counteract the fee reduction itself. Because revenue fees apply once a listing receives its first review, hosts can expect a drop in variable per-period profits as soon as the first review comes in. Raising their price reduces the chances of this happening. Hence, surprisingly, a simple fee reduction lowers hosts' incentives to accelerate social learning, or might even create an incentive to slow it down. Moreover, since the revenue fee waiver increase the variable per-period profits of new listings, hosts' entry incentives increase dramatically. The resulting increase in the number of listings together with the fact that prices barely decrease, lead to a decline in the occupancy rate of any individual new listing.

This suggests that reducing or even waiving Airbnb's revenue fee is not effective at increasing social learning and may even be counterproductive. In line with this conclusion, the decomposition in column (1) of Table 9 shows a large variety effect (reflecting higher entrant profits) which makes up more than 80% of the total effect, but a small but negative social learning effect due to lower occupancy rates for new listings. The price effect is positive as expected, but fairly small since the revenue fee reduction is barely passed on to consumers.

	(1)	(2)
	$r_0 = 0$	$r_0 = 0$ & $v_0 = 1$
Δ Total welfare	4.3%	5.2%
Δ Consumer welfare	6.0%	6.3%
Δ Total host profits	21.9%	9.0%
Δ Airbnb revenue	-8.2%	0.2%
Δ Entrant price	-0.8%	-0.3%
Δ Entrant occupancy rate	-1.0%	-1.7%
Δ # listings	6.8%	5.6%

Table 8: Effects of Airbnb (1) fully waiving the revenue fee only, and (2) waiving the revenue fee only if the guests leaves a review.

When the revenue fee reduction is tied to strong incentives for review provision by consumers, effectively garantueeing a review after each stay, the total welfare and consumer welfare effect are larger at 5.2% and 6.3% respectively (column (2) of Table 8). However, the increase in total host profits, with only 9%, is smaller relative to the increase of more than 21% in the previous counterfactual. Moreover, in contrast to the previous counterfactual, Airbnb's revenue slightly increases overall, despite the loss in revenue fees for entrant listings. Interestingly, the decline in entrant prices is even lower and the decline in entrant occupancy rate more pronounced than in the previous counterfactual. The total number of listings increases by 5.6%.

The reason for the smaller increase in total host profits is that, as the conditional review probability increases, new listing will benefit from the revenue fee reduction for a shorter duration. This is evidenced by the less pronounced increase in the total number of listings. Furthermore, a higher conditional review probability further incentivizes entrants to use price increases as a means to avoid losing the revenue fee waiver. Therefore, the entrant price reduction is now even smaller, and the decrease in entrant occupancy rates is more pronounced, underscoring the interplay between dynamic pricing incentives and consumers' probability of leaving a review. Nonetheless, guarantueed reviews after bookings appear to outweigh the counterproductive pricing response of sellers, resulting in an increase in total welfare, consumer welfare, and ultimately Airbnb revenue. The welfare decomposition in column (2) of Table 9 supports this interpretation – the variety effect is substantially smaller, while the social learning effect is the primary driver of the increase in consumer welfare.

Whereas consumer benefit in both counterfactuals, reducing the fee is only profitable for Airbnb if the conditional review probabilities increases. This suggests the existence of a set of combinations of revenue fee decreases and conditional review probability increases that Airbnb would find profitable to implement. Figure 10 depicts Airbnb revenue and consumer welfare for different combinations of these policy levers. The red line in both plots represents the iso-revenue line where Airbnb revenue is at status quo level. The interests of Airbnb and of the guests are aligned with respect to the conditional review probability as both Airbnb and guests benefit when the review rate is higher. However, Airbnb would prefer no revenue fee reduc-

	(1)	(2)
	$r_0 = 0$	$r_0 = 0$ & $v_0 = 1$
Δ Consumer welfare	6.05%	6.27%
Δ variety effect	5.72%	1.44%
Δ price effect	0.43%	0.21%
Δ social learning effect	-0.10%	4.63%

Table 9: Decomposition of the consumer welfare gain of Airbnb (1) fully waiving the revenue fee only, and (2) waiving the revenue fee under the condition of providing a truthful review.

tion, while consumers are better off with larger fee reductions. Airbnb is only willing to fully waive the revenue fee if it can condition the waiver on guests providing reviews with certainty. For conditional review probabilities below one, Airbnb may be willing to reduce, but not fully waive, the fee. The combination that maximizes consumer welfare among those profitable for Airbnb lies in the lower right corner of both panels in Figure 10, and is the one analyzed in column (2) of Table 8 and Table 9.

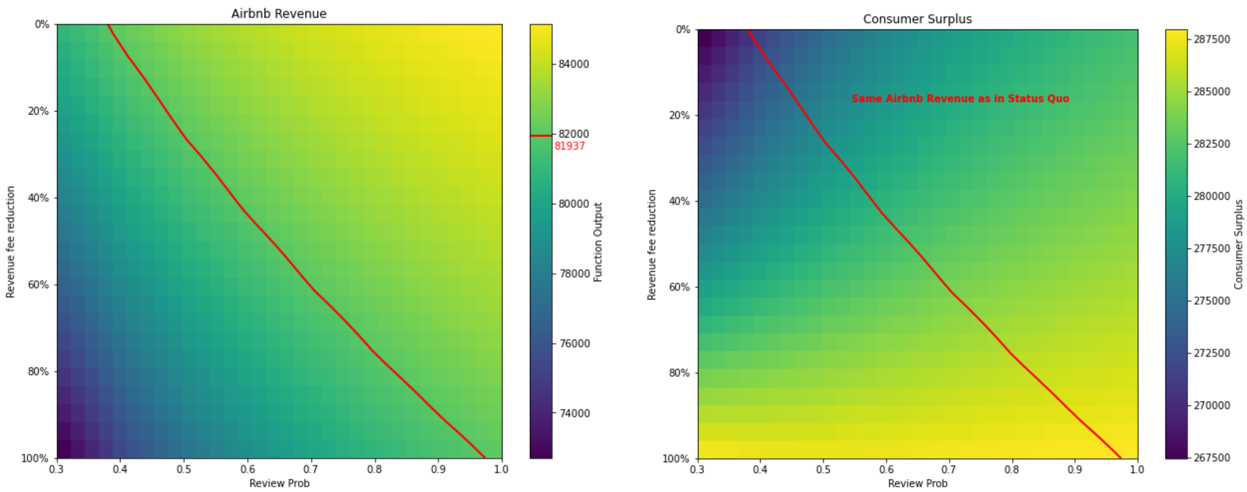

Figure 10: Airbnb revenue (left) and consumer surplus (right) for different levels of revenue fee reductions and review probabilities. The status quo review probability is 38% (for a zero revenue fee reduction).

In summary, the unconditional waiving of revenue fees for new listings counterintuitively leads to a slight *reduction in information* on these listings. While increasing the conditional review probability accelerates social learning, entrant hosts are incentivized to actively avoid gaining reviews as this would lead them to lose the higher profits under the revenue fee waiver. While the policy is aimed at accelerating social learning, it achieves the opposite due to hosts' misaligned pricing incentives. To have the intended effect on social learning, Airbnb has to tie the revenue fee waiver to review provision by guests. In the subsequent analysis, we explore a set of counterfactuals where Airbnb charges entrant hosts a lump-sum tax to finance the revenue fee reduction. This lowers the profitability of new listings, as it is equivalent to forcing hosts to reduce their prices.

9.2.2 Hosts pay for the revenue fee reduction

Table 10 summarizes our results for two cases: (1) Airbnb waves the revenue fee but compensates the revenue losses by charging entrants a lump-sum tax, (2) Airbnb uses the policy in (1) to incentivize consumers to provide reviews (similar to the Vine program on Amazon). In both cases, the lump-sum tax is equal to the original revenue fee of 14.2% times the type-specific expectation over entrant revenue in the counterfactual.

Waiving the revenue fee now increases total welfare by 0.5% (column (1) in Table 10). While this also leads to increases in consumer welfare and Airbnb revenue, total host profits decline by 4%. Accordingly, entry incentives are smaller, resulting in a 1.7% decrease in the total number of listings. The revenue fee waiver effectively reduces the price consumers pay by 5%, which in turn leads to a 6% increase in occupancy rates.

The increase in entrant occupancy rates contrasts with the first counterfactual in the previous section, where the occupancy rate slightly decreased. This difference comes down to two factors. First, with demand in effect spread across fewer listings, the occupancy rate of the remaining listing increases. Second, and more importantly, because Airbnb's revenue fee loss is now passed on to new listings, these hosts have an added incentive to obtain reviews in order to avoid the lump-sum tax. In other words, their dynamic pricing incentives are aligned with the policy objective of increasing the speed of social learning, with substantial price reductions faced by consumers and, hence, occupancy rate increases. The welfare decomposition results in column (1) of Table 11 underlines the fact that Airbnb hosts' endogenous decision support the policy objective now. While the variety effect is negative, it is smaller than the sum of the (positive) price and social learning effects. The social learning effect is positive and accounts for around 40% of the total gain in consumer welfare.

	(1)	(2)
	$r_0 = 0$	$r_0 = 0$ & $v_0 = 1$
Δ Total welfare	0.5%	3.8%
Δ Consumer welfare	0.9%	4.4%
Δ Total host profits	-4.0%	-2.8%
Δ Airbnb revenue	1.1%	4.4%
Δ Entrant price	-5.0%	-5.3%
Δ Entrant occupancy rate	6.0%	6.4%
Δ # listings	-1.7%	2.2%

Table 10: Effects of Airbnb (1) fully waiving the revenue fee in exchange for a revenue-neutral lump-sum tax on entrants, (2) waiving the revenue fee in exchange for a revenue-neutral lump-sum tax on entrants under the condition of providing a truthful review.

When consumers provide reviews in return for the revenue fee being waived, the overall welfare effect is larger with 4.4%, as are the gains in consumer welfare and Airbnb revenue (column (2) in Table 10). Additionally, the decrease in total host profits is less pronounced. At the same time, the decrease in entrant price is larger, as is the increase in the occupancy rate of entrants. In contrast to the previous counterfactual, the total number of listings increases, by 2.2%.

The reason for the smaller loss in total host is that a higher conditional review probability increases the probability of gaining a review and avoiding the cost of the lump-sum tax associated with revenue fee waiver. The higher conditional review probability also interacts with the dynamic pricing incentives, but in the opposite direction compared to the set of counterfactuals considered in Section 9.2.1. The probability of gaining a review is now more elastic with respect to prices, providing entrant listings an additional incentive to reduce them. This leads to a greater increase in the occupancy rate of entrants. Interestingly, the overall number of listings increases, despite the reduction in variable per-period profits for new listings. This suggests that faster social learning is the primary driver of the increase in the number of listings. The consumer welfare decomposition in column (2) of Table 11 corroborates this. While the variety effect remains negative, meaning the change in variable per-period profits alone would have resulted in fewer listings, the large, positive social learning effect significantly outweighs the negative variety effect.

	(1) $r_0 = 0$	(2) $r_0 = 0$ & $v_0 = 1$
Δ Consumer welfare	0.86%	4.42%
Δ variety effect	-1.75%	-1.71%
Δ price effect	2.26%	1.19%
Δ social learning effect	0.35%	4.93%

Table 11: Decomposition of the consumer welfare gain of Airbnb (1) fully waiving the revenue fee only, (2) waiving the revenue fee under the condition of providing a truthful review.

We find that there is no combination of entrant-financed revenue fee reduction and review probability increase that would make hosts overall better off. While increasing the rate at which new listings receive their first review benefits new listings, it harms incumbent listings for two reasons: First, the elasticity of the unconditional review probability with respect to price increases, incentivizing new listings to lower their prices, causing competition to intensify. Second, new listings are reviewed more rapidly, leading to a larger number of incumbents and further increasing competitive pressure. The profit reductions due to this heightened competitive pressure on incumbents appear to outweigh the faster accumulation of the first review.

Compared to the counterfactuals where Airbnb absorbed the revenue fee waiver in Section 9.2.1, when entrants bear the cost through a lump-sum tax, the welfare gains from the policy are smaller in general. Charging a lump-sum tax to cover the revenue fee waiver diminishes their profits and entry incentives while increasing their exit incentives. As a result, the negative variety effect must be compensated by larger price and social learning effects. It turns out that both the price and social learning effects are indeed larger than in the previous section's counterfactuals. In particular, prices decrease substantially. This price decrease, combined with the slight reduction in the number of listings, leads to higher occupancy rates for entrant listings and therefore faster information diffusion. Consequently, the social learning effect is a significant component of the overall increase in consumer welfare.

10 Conclusion

In this study, we have examined the implications of different policy interventions by Airbnb aimed at accelerating social learning and information diffusion on its platform, taking into account interactions with endgenous supply-side decisions, such as pricing, entry and exit. 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 estimates indicate that reviews significantly influence guests' booking decisions and host profits.

We analyzed the welfare and market outcomes of policies that waive the revenue fee for new entrant listings, either unconditionally or by financing the waiver through a lump-sum tax charged to entrants. Our findings demonstrate that the revenue fee waiver policies can indeed increase the speed of social learning, but the effects are mediated by complex interactions between pricing incentives, entry/exit dynamics, and the review process. An unconditional revenue fee waiver counterintuitively leads to less information accumulation on new listings, as entrants seek to avoid the loss of higher profits associated with gaining reviews.

Alternatively, financing the revenue fee waiver through a lump-sum charge to entrants aligns their dynamic pricing incentives more closely with the goal of accelerating social learning. This policy reduces entrant profitability, incentivizing them to more actively seek reviews in order to avoid the lump-sum tax. As a result, this approach is more effective at enhancing information diffusion, with the social learning effect making up a significant portion of the overall welfare gains.

These results underscore the importance of carefully considering the complex, interconnected incentives facing different platform participants when designing policies intended to shape information aggregation and diffusion. A nuanced, multi-faceted approach that aligns these incentives appears necessary to fully harness the potential welfare gains from accelerated social learning on two-sided platforms like Airbnb.

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.

AirDNA, "Airbnbs in New York, Newark and Jersey City," 2019.

Almagro, Milena and Tomás Domínguez-Iino, "Location sorting and endogenous amenities: Evidence from amsterdam," *Econometrica*, 2025, *93* (3), 1031–1071.

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.

Bolton, Patrick and Christopher Harris, "Strategic experimentation," *Econometrica*, 1999, *67* (2), 349–374.

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

Butler, Jeffrey V., Enrica Carbone, Pierluigi Conzo, and Giancarlo Spagnolo, "Past performance and entry in procurement: An experimental investigation," *Journal of Economic Behavior Organization*, 2020, *173*, 179–195.

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.

Conlon, Christopher T and Julie Holland Mortimer, "Demand estimation under incomplete product availability," *American Economic Journal: Microeconomics*, 2013, *5* (4), 1–30.

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

Doraszelski, Ulrich 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.

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.

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.

Kovbasyuk, Sergey and Giancarlo Spagnolo, "Memory and Markets," *The Review of Economic Studies*, 06 2023, *91* (3), 1775–1806.

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*, 2024.

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, 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.

Appendix

A.1 Simple Model with endogenous prices

The expected indirect utilities of product $j \in \{E, I\}$ are given in Equation (20). The unit price of product j is denoted by p_j and ϵ_{jt} represents the idiosyncratic taste shock.

$$u_{ijt} = \omega_{jt} - p_{jt} + \epsilon_{jt} = \nu_{jt} + \epsilon_{ijt}, \;\text{ where }\; \epsilon_{ijt} \overset{iid}{\sim} Gumbel(0, \pi^2/6) \tag{20}$$

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 (21). 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 $t = 1$.

$$\max_{p_{j1}} \left(q_{j1} p_{j1} + \delta \mathbb{E}_2 \left[q_{j2} p_{j2} | p_{E1}, p_{I1} \right] \right) \tag{21}$$

Equation (21) 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 A.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*

$$p^*_{E1} = \frac{1}{1 - q_{E1}(p^*_{E1}, p^*_{I1})} - v_r \delta \sum_{i \in \{0,g,b\}} \phi_i \pi^i_{E2}$$
$$\textit{and}\;\; p^*_{I1} = \frac{1}{q_{E1}(p^*_{E1}, p^*_{I1})} + v_r \delta \sum_{i \in \{0,g,b\}} \phi_i \pi^i_{I2}.$$

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

Proof. 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}.

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

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.

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

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 $\partial p^*_{I1}/\partial \delta > 0$. □

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 A.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 A.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 A.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 $\tilde{p}^s_1 = p^s_{E1} - p^s_{I1}$.[29]

Proposition A.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.,

$$\tilde{p}^s_1 < 0 \qquad \forall \qquad \tilde{\omega}_1.$$

(ii) If $\sum_i \delta v_r \phi_i(\pi^i_{I2} - \pi^i_{E2}) < 1$, there exists $\widehat{\omega_1} < 0$ such that $\tilde{p}^s_1 < \tilde{p}^*_1 \quad \forall \tilde{\omega}_1 > \widehat{\omega_1}$.

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

Proof. **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 $(\omega_I - \tilde{p}_I)\ln\left(1 + \exp\left(\tilde{\omega}_2 - \tilde{p}_2\right)\right) + \text{constant}$ 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}:

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

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

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

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 $\tilde{p}_1^s$.

$$\tilde{p}_1^s = -\sum_{i \in \{0,g,b\}} v_r \delta \phi_i \left(u_2^i + \pi_{E2}^i + \pi_{I2}^i\right) \tag{22}$$

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 Appendix A.1 and Equation (22), we know:

$$\begin{aligned}
\tilde{p}_1^s < \tilde{p}_1^* \iff -\sum_i v_r \delta \phi_i u_2^i &< \frac{1}{1 - q_{E1}^*(\tilde{\omega}_1)} - \frac{1}{q_{E1}^*(\tilde{\omega}_1)} \\
&= \exp\left(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)\right) - \frac{1}{\exp\left(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)\right)}
\end{aligned} \tag{23}$$

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 Appendix A.1 for entrant and incumbent:

$$\tilde{p}_1^*(\tilde{\omega}_1) = \exp\left(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)\right) - \frac{1}{\exp\left(\tilde{\omega}_1 - \tilde{p}_1^*(\tilde{\omega}_1)\right)} - \sum_i \delta v_r \phi_i (\pi_{E2}^i + \pi_{I2}^i) \tag{24}$$

Notice that for any ω_{E1}, Equation (23) implies that, if $\tilde{\omega}_1 = 0$, $\tilde{p}_1^* < 0$. Therefore,

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

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 (23) is decreasing as ω_{I1} decreases or $\tilde{\omega}_1$ increases.

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

As long as $-1 < \frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}}$, the RHS of inequality (23) increases as ω_{I1} decreases or $\tilde{\omega}_1$ increases.

Using the Implicit Function Theorem on Equation (24), we can derive $\frac{d\tilde{p}_1}{d\tilde{\omega}_{1I}}$:

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

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

□

According to Proposition A.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 A.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.

We simulate the model with $\omega_I = 1$ for various values of a and b, covering the range $\frac{1}{11} \leq \omega_{E1} \leq \frac{10}{11}$. For the purpose of the model simulation, we set the price and quality coefficients to 0.1 and 2 respectively instead of 1. Figure A.1 illustrates the variation in $\tilde{p}_1^S - \tilde{p}_1^*$ as

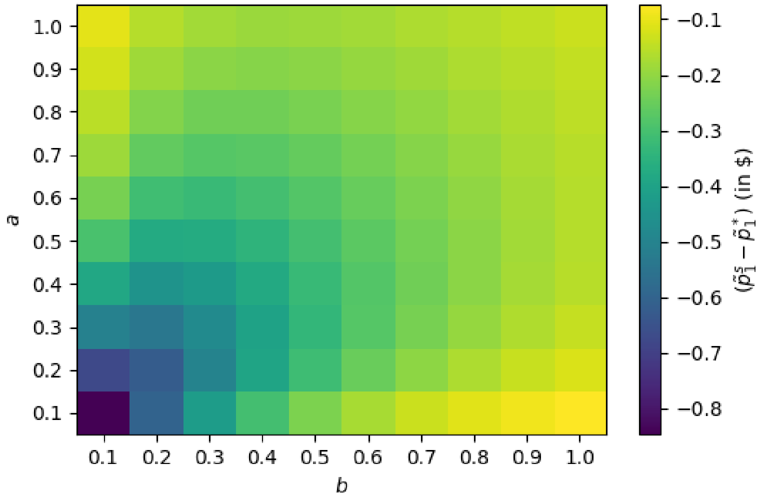

Figure A.1: Model simulations for different ω_{E1} ($\omega_I = 0.5$, the price coefficient is 0.1, and there are 1,000 consumers).

a function of the expected quality of the entrant's product. The cold-start problem arises when $p_1^S - p_1^* < 0$ and as more severe as $p_1^S - p_1^*$ decreases, represented by 'colder' colors in the heat map. Consistent with Proposition A.1, we observe that the cold-start problem occurs when consumers' prior is uncertain and the social value of the review is high. Then, the social planner sets a particularly low price, compared to the entrant, to encourage experimentation.

Proposition A.1 stands in contrast to the existing literature. In the absence of endogenous prices, informational externalities have been shown to give rise to the cold-start problem (Bolton and Harris, 1999). However, with endogenous pricing, Bergemann and Välimäki (2000) show that price competition can lead to *over*exploration. In their model, price competition softens as consumers uncover the quality of a new product – a dynamic individual consumers fail to consider. Consumers derive no benefit from exploration because products are horizontally undifferentiated, and the consumer surplus function is linear in product quality. In contrast, with logit demand consumers benefit from discovering the product quality of the entrant's product. As a result, the cold-start problem can arise in our model, even when firms set prices.

Online Appendix

B.1 First-order derivative of the pricing stage

Remember that hosts set their prices after they learn about their marginal cost shock, but they do not know about their rival's cost shock. To decrease the computational burden, we assume that hosts react to average prices by competitors instead of the full price distribution induces by $P_0(x, c)$. Hence, the FOC is given by:

$$w'(x, c|P_1^k, P_0) = \mathbb{E}\left[q(P_1^k, \mathbb{E}_0[p_{-j}], x_j) + \frac{\partial q(P_1^k, \mathbb{E}_0[p_{-j}], x)}{\partial p}(P_1^k - c) + \delta\frac{\partial \mathbb{E}_{x'}[V(x') - (1-\chi(x'))\bar{\phi}|\mathbb{E}_0[p_{-j}]]}{\partial P_1^k}\right]$$

where $\mathbb{E}_0[p_{-j}] = \mathbb{E}_c[p_{-j}|p_{-j} = P_0(x, c)]$.

B.2 Expected operating cost

The expected operating cost is given by the following expression.

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

B.3 Expected entry cost

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

$$\begin{aligned}\lambda\mathbf{E}_\kappa[\kappa|\kappa \leq \delta\mathbb{E}_l[V((0,0,l))]] =& \bar{\kappa} - (1-\lambda)\mathbb{E}_\kappa[\kappa|\kappa > \delta\mathbb{E}_l[V((0,0,l))]]\\ =& \bar{\kappa} - (1-\lambda)\left(\kappa + \delta\mathbb{E}_l[V((0,0,l))]\right)\\ =& \lambda\bar{\kappa} - (1-\lambda)\delta\mathbb{E}_l[V((0,0,l))]\end{aligned} \tag{25}$$

B.4 Additional Tables and Figures

	Dependent variable:			
	p	N	K	s_g
	(1)	(2)	(3)	(4)
reservation length	-0.829***	-4.262***	-3.944***	-0.00000
	(0.162)	(0.094)	(0.087)	(0.00000)
rating	17.955***	-0.208	2.339***	0.0002***
	(1.502)	(0.542)	(0.438)	(0.00001)
$B_{t-3} - \bar{B}$	-2.199	-5.318	-5.098	-0.00000
	(21.781)	(7.676)	(6.042)	(0.0002)
$B_{t-4} - \bar{B}$	-2.202	-1.134	-3.624	0.001**
	(23.729)	(8.111)	(6.454)	(0.0002)
$B_{t-5} - \bar{B}$	-6.246	1.540	-5.286	-0.001**
	(18.629)	(7.890)	(6.338)	(0.0002)
# Airbnb listings	0.013***	-0.013***	-0.013***	-0.00000***
	(0.003)	(0.002)	(0.002)	(0.00000)
rating $\times(B_{t-3} - \bar{B})$	0.008	1.268	1.209	0.00001
	(4.600)	(1.611)	(1.275)	(0.00005)
rating $\times(B_{t-4} - \bar{B})$	0.547	0.229	0.739	-0.000**
	(5.006)	(1.703)	(1.363)	(0.00005)
rating $\times(B_{t-5} - \bar{B})$	2.622	1.967	3.249**	0.0001*
	(3.937)	(1.658)	(1.338)	(0.00004)
Constant	108.552***	74.559***	57.930***	0.002***
	(7.707)	(3.214)	(2.677)	(0.0001)
Observations	24,816	24,816	24,816	24,816
R^2	0.018	0.065	0.066	0.060
Adjusted R^2	0.018	0.064	0.066	0.059
Residual Std. Error (df = 24806)	58.921	41.549	38.459	0.001
F Statistic (df = 9; 24806)	51.611***	190.523***	195.528***	174.937***

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

Table B.1: Demand estimation: first stage regression results

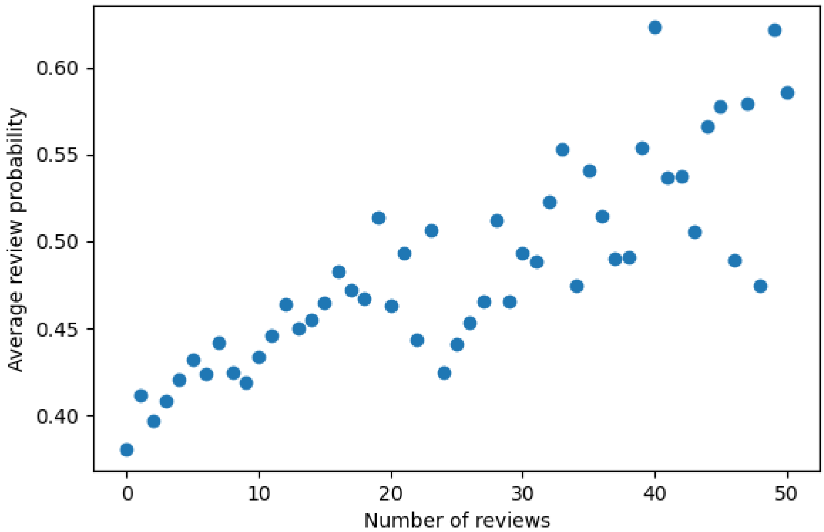

Figure B.1: Empirical review probabilities conditional on the number of reviews.