Value for Money and Selection: How Pricing Affects Airbnb Ratings *

Christoph Carnehl[†] Maximilian Schaefer[‡]

André Stenzel[§] Kevin Ducbao Tran[¶]

March 20, 2025

Abstract

We study strategic pricing incentives when prices affect ratings. Two channels determine the price-rating interaction: higher prices reduce the value for money, worsening ratings, but increase the taste-based valuation of travelers, improving ratings. Our empirical results in the context of Airbnb show a dominant value-for-money effect and that hosts benefit from low entry prices: offering a median entry discount of seven percent improves medium-run monthly revenues by three percent. We also provide some evidence in favor of strategic host behavior. More professional hosts are more likely to use entry discounts, and hosts lower their prices to surpass rating thresholds.

JEL Codes: D18, D25, D47, D82
Keywords: Rating Systems, Dynamic Pricing, Asymmetric Information

*We thank seminar participants at the University of Bologna, the University of Bristol, the University of Leicester, Northwestern University, Yale University, and the JRC Seville, the Mannheim Economics Alumni Symposium, 2023 Retail Payments, Multi-Sided Markets and Platforms Workshop at the Bank of Canada, the 2023 CEA, the MaCCI 2021 Annual Conference, the Louvain and Paris Economics of Digitization Online Seminar, the 13th Paris Conference on Digital Economics, the 11th bi-annual Postal Economics Conference at Toulouse School of Economics, the 13th EIEF-IGIER-UNIBO IO workshop, the 2024 VfS Industrial Economics Committee Meeting, the 2024 Lisbon Meetings in Game Theory and Applications, the 2024 IBEO Workshop Digital Economics meeting, EARIE 2024, as well as Jason Allen, Long Bui, Laura Grigolon, Johannes Johnen, Robin Ng, Matthew Osborne, Helena Perrone, Imke Reimers, Michelangelo Rossi, Konrad Stahl, Hannes Ullrich, and Christine Zulehner for insightful comments. We are grateful to Glenna Jenkins for editorial assistance. This work would not have been possible without the InsideAirbnb project, which provided an essential part of the data free-of-charge. All errors are our own.

[†]Bocconi University, Department of Economics and IGIER, christoph.carnehl@unibocconi.it

[‡]Institut Mines-Télécom Business School, maximilian.schafer@imt-bs.eu

[§]Bank of Canada, astenzel@bankofcanada.ca. The views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank. Funding by the German Research Foundation via CRC TR 224, Project C03, during prior employment at the University of Mannheim is gratefully acknowledged.

[¶]University of Bristol, kevin.tran@bristol.ac.uk

1 Introduction

Online exchanges such as Airbnb, eBay, and Deliveroo typically match buyers and sellers who have not previously interacted. As the traded goods or services cannot be physically examined pre-purchase, these platforms seek to build and foster trust between participants to resolve inherent quality uncertainty. A key instrument for this purpose is the use of reputation and feedback systems, which are prevalent on virtually every online exchange (see Tadelis, 2016). These systems aim to provide agents with reliable signals about the quality of the other side of the market. Therefore, understanding how ratings are generated and whether agents can strategically influence them is essential to designing effective reputation and feedback systems.

In this paper, we derive and empirically assess predictions regarding pricing incentives when prices affect individual review scores and aggregate ratings.[1] We derive predictions from an illustrative theoretical model and utilize a unique transactions and ratings dataset of Airbnb listings in Paris, France, for 2017. Our data allow us to match booking prices to the evolution of aggregate ratings. First, we show that, all else equal, higher prices are associated with lower review scores. Second, consistent with our theoretical analysis, we provide evidence that newly entering hosts benefit in the medium run from offering early discounts. An initial discount of five euros per night (seven percent) increases medium-run monthly revenues by approximately 68 euros (three percent) compared to listings that do not offer a discount when entering the market.

In our model, a long-lived host offers an apartment of fixed quality to short-lived travelers. In each period, the host chooses the listing price. Prospective travelers observe an aggregate rating—which summarizes individual reviews—and use this to form beliefs about the quality of the apartment. The booking decisions of travelers depend on these beliefs, their idiosyncratic preferences for the listing, and its price. Travelers who stay at an apartment may provide a review that depends on the apartment's actual quality, the travelers' idiosyncratic tastes, and the price. We assume that higher quality and taste induce higher review scores, and that higher prices directly lower review scores—we call the latter the *value-for-money effect*. However, higher prices also induce a *selection effect*: a price increase induces the previously marginal traveler to forgo booking and thus increases the idiosyncratic taste of the average traveler who stays at the apartment. As a result, higher prices indirectly lead to higher review scores.

The net effect of prices on the induced review scores, i.e., the sign of the *induced review effect*, depends on the relative importance of the value-for-money and selection effects. Importantly, the host's current pricing decision will influence the reviews posted by current-period travelers, thereby the aggregate rating observed by future consumers and future profits. We use the model to shed light on the host's incentives and to derive testable hypotheses on the relationship between prices and ratings.

On Airbnb, travelers can leave an overall review, as well as a review across six subcategories: value-for-money, location, accuracy of the listing description, cleanliness,

[1] Throughout the paper, we refer to the evaluation left by an individual consumer as a *review*, while a *rating* is an aggregate measure summarizing reviews by multiple consumers.

host-communication, and the check-in experience. Our first set of empirical results indicates that, in the context of our data, the value-for-money effect dominates the selection effect. When regressing review scores on lagged prices and controlling for listing-specific fixed effects, we find that higher prices are generally associated with lower scores. However, the strength of the relationship differs across rating categories: it is particularly pronounced for value-for-money ratings and disappears for location ratings. While the overall negative relationship between prices and ratings suggests that the value-for-money effect dominates the selection effect, the absence of this relationship for location ratings is in line with a relatively stronger selection effect for this specific rating category. This result is sensible because the location rating asks the traveler about a component specifically related to taste.

In light of our model, the results indicate that strategic hosts have an incentive to charge lower prices, all else equal, to boost their ratings. In the remainder of our empirical analysis, we explore two additional hypotheses that derive from the dominant value for money effect. First, when analyzing hosts' pricing behavior in relation to salient rating thresholds, we find evidence that some hosts appear to be aware of and strategically exploit the dominant value-for-money effect.[2] Using variation in how close hosts are to thresholds that discontinuously improve their aggregate star rating, we find that hosts closer to a threshold tend to lower their prices relative to those further away. This result is intuitive in light of our model because it reflects the dominant value-for-money effect which incentivizes strategic hosts to charge lower prices to gain a rating advantage. Second, our theoretical analysis suggests that entry price discounts, which sacrifice short-run profits, can increase medium-run profits by persistently improving the ratings. The underlying reason is that prices can be adjusted frequently, but they have a persistent effect on ratings, which determine future demand. Given that the value-for-money effect dominates, hosts should offer a price discount when entering the market relative to a naive entry price that ignores this dynamic effect.

Our theoretical hypothesis with respect to entry price discounts is supported by our empirical analysis focusing on listings that enter the market. We find that listings that enter the market with a discount receive better ratings and more bookings in the early periods. This allows these hosts to set relatively higher prices in subsequent periods without an associated quantity reduction, resulting in higher revenues in the medium run. Specifically, a discount of five euros (seven percent) when entering increases medium-run monthly revenues by approximately 68 euros (three percent). Finally, we provide separate evidence suggesting that hosts who offer an entry discount tend to be more professional. Our interpretation of this observation is that professional hosts tend to be more strategic in their decision and thus factor the dominant value-for-money effect in when setting entry prices.

Related literature. The literature on the ratings-prices nexus has mostly focused on how ratings affect prices. Several studies establish a robust positive relationship

[2]The impact of salient rating thresholds on consumers has been studied in the empirical literature starting with the seminal work by Luca (2016).

between ratings and prices (Teubner et al., 2017), revenues (Luca, 2016; Fang, 2022), and quantities (Livingston, 2005).

In contrast, we are interested in whether and how prices affect ratings, contributing to the literature on strategic ratings management through price setting. While our main contribution is empirical, our theoretical model is closely related to Carnehl et al. (2024), albeit with a different focus. They study the theoretical long-run properties of learning, ratings, and prices, as well as the design of the rating system. Our model emphasizes the effects at play and, in addition, explicitly accounts for whether a review is generated in a given period. Moreover, we focus on the effects of prices set in the early stages of a listing's life cycle. Our work is thus related to Johnen and Ng (2024), who incorporate the value-for-money effect into a model of signaling through prices and ratings. They show that sellers harvest high ratings obtained from consumers' reciprocal response to low initial prices.[3] Overall, we contribute to this literature by empirically assessing the net effect of prices on ratings for a specific platform market, and by quantifying the benefits of engaging in strategic price setting upon entry.

Empirically, Zegners (2019) finds that books offered for free on an online self-publishing platform generate more but worse ratings. In line with the selection effect in our model, the author argues that this result is because readers who read a free book have a lower preference for it. We add to this line of research by considering continuous variation in prices rather than comparing a price of zero and positive prices.[4] Furthermore, we consider an additional effect that prices can have on ratings: the value-for-money effect. Luca and Reshef (2021) analyze daily menu prices and ratings on an online ordering platform and find that price increases lead to decreases in average ratings. Relatedly, Sorokin (2021) finds that producers on the video game platform Steam use discounts to transition to higher ratings tiers, while Jeziorski and Michelidaki (2023), who analyze different sub-categories of ratings, empirically show that the explicit value-for-money rating features the highest degree of price responsiveness. Our theoretical framework can accommodate these seemingly conflicting empirical results, and our empirical evidence is in line with a dominant value-for-money effect in the context of Airbnb.

Our paper is more broadly related to research on the determinants of ratings. Cabral and Li (2015) find that lower quality transactions result in more negative feedback. Proserpio and Zervas (2017) find that when responding to consumer feedback, hotels tend to receive fewer but longer negative reviews. Rossi and Schleef (2024) finds that movie ratings tend to be lower after a movie was nominated for an Academy Award, in line with consumer disappointment. In addition to legal determinants, there is ample research on the use of fake reviews to improve ratings, see Mayzlin et al. (2014); Luca and Zervas (2016); He et al. (2022). We add to this literature by highlighting theoretically and empirically how price setting—as a specific

[3]Our work is also related to Bondi (2025) who models consumer learning from reviews in a setting with consumer self-selection and quality uncertainty. The main difference compared to our setting is that the price does not directly affect ratings.

[4]This difference is particularly relevant in light of research that finds that a zero price can have a differential effect on demand (Shampanier et al., 2007).

strategic action by sellers—can affect ratings.

This also sets our study apart from other empirical analyses focusing on the Airbnb platform beginning with Fradkin et al. (2021), which, e.g., focus on the competitive effects of changes in the cancellation policy (Jia et al., 2021) and short-term rental regulation (Rossi, 2024), the welfare effects of peer entry (Farronato and Fradkin, 2022) or on safety reviews (Culotta et al., 2023). By quantifying the impact of prices on reviews and ratings, our analysis complements recent work by Huang (2022) and Foroughifar and Mehta (2023), who study algorithmic pricing tools on the platform. The key difference is that we do not focus on the source of sellers' pricing decisions but instead on how these decisions affect induced review scores and, thereby, future profits.

We proceed as follows. In Section 2, we introduce and study our theoretical framework. In Section 3, we introduce the data and provide descriptive statistics. We describe our empirical analysis and results of the price-rating interactions in Section 4. In Section 5, we empirically quantify the value of entry-price discounts. We discuss implications for hosts and the platform operator in Section 6 and conclude in Section 7.

2 Theory: Prices, Ratings & Dynamic Incentives

We begin by setting up and analyzing a stylized theoretical framework to guide our empirical analysis. Our model consists of a host offering an apartment to a sequence of potential travelers. We outline each model component in detail below.

Host. There is a single long-lived host with an apartment of fixed quality, $\theta \in \{L, H\}$.[5] The common prior probability that the good is of high quality is given by $\mu_0 \in (0, 1)$. In each period $t = 1, \ldots, T$ with $T \leq \infty$, the host chooses a price $p_t \geq 0$. The host has a discount factor of $\delta \in [0, 1)$ and maximizes their discounted sum of profits.

Travelers. In each period, a potential traveler arrives. Travelers value the quality of the apartment, their idiosyncratic taste for the apartment, and money. The idiosyncratic taste ω_i is uniformly distributed, $\omega_i \sim U[0, 1]$. At the time of purchase, travelers do not know the quality of the apartment. However, they have access to an aggregate rating $\bar{\Psi}_t$, which they observe before deciding whether to book.

Reviews and Ratings. The aggregate rating $\bar{\Psi}_t$ is the star rating displayed when travelers search for potential apartments, ranging from one to five stars in half-star increments. We reflect this in our model setup by having $\bar{\Psi}_t$ obtained by rounding

[5]While the empirical setting naturally features multiple competing hosts, the restriction to a monopoly is for expositional purposes only—the core mechanisms driving our hypotheses are not affected by accounting for additional hosts. For further details, see Carnehl et al. (2024), who analyze extensions to competitive settings in a related model.

the latent average

$$\lambda_t = \sum_{s=1}^{t-1} \frac{\mathbb{I}_{\text{review left in } s} \cdot \psi_s}{\sum_{s=1}^{t-1} \mathbb{I}_{\text{review left in } s}} \tag{1}$$

of past review scores ψ_s to the nearest half-star. This simple averaging is in line with the aggregate rating formation on Airbnb; see Appendix B. It also reflects the most prominent information about past consumers' reviews when searching on Airbnb in 2017, i.e., when our data were collected. Note that Airbnb has since changed the display of search results to contain a non-rounded average rating score. For simplicity, we assume that every traveler leaves a review with an identical probability ρ.[6]

We impose the following assumptions on the formation of individual reviews ψ_i, which reflect the quality of the apartment θ, the idiosyncratic traveler taste ω_i and the price p, and are also on the scale from one star to five stars.

Assumption 1 *Individual review scores are in expectation*

(i) increasing in apartment quality, $E[\psi_i|H, \omega_i, p] > E[\psi_i|L, \omega_i, p]$*,*

(ii) increasing in the traveler's idiosyncratic taste, $\frac{\partial}{\partial \omega} E[\psi_i|\theta, \omega_i, p] > 0$*, and*

(iii) decreasing in the booking price of the apartment $\frac{\partial}{\partial p} E[\psi_i|\theta, \omega_i, p] < 0$*.*

We impose the assumptions in expectation only to allow for potential randomness in consumer reviews. Given the averaging of past reviews and denoting the *realized* review in a given period (i.e., if the apartment was booked and a review was left) by ψ_t, we obtain the law of motion of the latent aggregate rating

$$\lambda_{t+1} = \frac{t}{t+1} \lambda_t + \frac{1}{t+1} \psi_t, \tag{2}$$

with $\lambda_{t+1} = \lambda_t$ if no review was left and the salient star rating $\bar{\Psi}_{t+1}$ obtained by rounding λ_{t+1} to the nearest half-star.

Belief formation and demand. Based on the observed rating, travelers form a belief about the apartment's quality. To illustrate the main trade-off faced by hosts, we model this belief formation process in reduced form with the assumption that consumer beliefs only depend on the salient aggregate rating and are increasing in this rating.[7] We denote the belief that the apartment is of high quality by $\mu(\bar{\Psi}) \in [0, 1]$, so that $\mu(\bar{\Psi}_1) > \mu(\bar{\Psi}_2) \iff \bar{\Psi}_1 > \bar{\Psi}_2$. This assumption reflects that travelers

[6] Relaxing this assumption does not alter the qualitative trade-offs we subsequently illustrate. Note that our setup independently features stochastic review generation because whether a booking takes place in a given period depends on that period's traveler's idiosyncratic taste.

[7] Having consumer beliefs only depend on the aggregate rating allows us to cleanly illustrate the mechanisms at play. In practice, they naturally depend on additional considerations, such as the number of reviews already posted, and we refer to Carnehl et al. (2024) for a detailed discussion of how they impact pricing incentives.

associate a higher rating—all else equal—with a higher quality. This feature emerges endogenously from the process by which individual reviews are generated, irrespective of the specific micro-foundation of $\mu(\cdot)$.

For ease of exposition, we assume that travelers are risk neutral and have an additively separable utility function resulting in the following expected utility, given their beliefs and the posted price p: $U(\bar{\Psi}, \omega_i, p) = \mu(\bar{\Psi}) + \omega_i - p$. Normalizing the travelers' outside option to zero, it follows that a traveler will book the apartment if and only if $U(\bar{\Psi}, \omega_i, p) \geq 0$. Therefore, there is an indifferent traveler with cutoff taste $\tilde{\omega} = p - \mu(\bar{\Psi})$ such that all travelers with $\omega_i \geq \tilde{\omega}$ would book the apartment.

Given that the taste parameters are uniformly distributed, we obtain the following likelihood that the apartment is booked in a given period, which we denote by q: $q(\bar{\Psi}, p) = 1 + \mu(\bar{\Psi}) - p$.[8]

2.1 Pricing Incentives

In this part, we consider the implications of the interaction between prices and ratings in our model for a host's pricing incentives. Our objective is to generate simple theoretical insights guiding the interpretation of the empirical price-rating interaction and the dynamic patterns observed in the data to quantify the potential benefits of strategic pricing.

Myopic host. As a benchmark, consider a myopic host who maximizes flow profits without considering the effect of the current price on future ratings. Such a host sets the myopic monopoly price given travelers' beliefs:

$$p_t^m = \frac{1 + \mu(\bar{\Psi}_t)}{2}. \tag{3}$$

Strategic host. In contrast, a strategic host takes the effect of the current price on future profits into account. Thus, the host maximizes the discounted sum of flow profits,

$$\max_{(p_t)_{t=1}^T} \sum_{t=1}^{T} \delta^t \left((1 + \mu(\bar{\Psi}_t) - p_t) p_t \right), \tag{4}$$

subject to the law of motion of the ratings defined above. We rewrite this problem in terms of the corresponding Bellman equation to obtain

$$V_t(\lambda_t) = \max_{p_t} \left(1 + \mu(\bar{\Psi}_t) - p_t) p_t + \delta E \left[V_{t+1} \left(\lambda_{t+1}(p_t) \right) \right] \right). \tag{5}$$

[8]We implicitly assume $\mu(\bar{\Psi}) \leq p \leq 1 + \mu(\bar{\Psi})$ such that this probability is well-behaved. This restriction is endogenously satisfied given the optimal pricing strategy of our host.

Taking into account the probability of a booking conditional on the price, as well as the updating rule, this can be further decomposed into

$$
\begin{aligned}
V_t(\lambda_t) = \max_{p_t} \quad \Bigg\{ & \underbrace{(1+\mu(\bar{\Psi}_t)-p_t)p_t}_{\text{flow profits}} \\
& + \Bigg(\underbrace{p_t - \mu(\bar{\Psi}_t)}_{Pr\{\text{no sale}|p_t\}} + \underbrace{(1+\mu(\bar{\Psi}_t)-p_t)\cdot(1-\rho)}_{Pr\{\text{sale and no review}|p_t\}} \Bigg) \cdot \delta \cdot V_{t+1}(\lambda_t) \qquad (6) \\
& + \underbrace{(1+\mu(\bar{\Psi}_t)-p_t)\cdot\rho}_{Pr\{\text{sale and review}|p_t\}} \cdot \delta \cdot E\left[V_{t+1}\left(\frac{t}{t+1}\lambda_t + \frac{1}{t+1}\psi_i\right) \middle| \, \omega_i \geq p_t - \mu(\bar{\Psi}_t) \right] \Bigg\}.
\end{aligned}
$$

Note that the state variable is the latent aggregate rating λ_t, which maps into the salient aggregate rating $\bar{\Psi}_t$ that consumers use to form beliefs. The expectation in (6) is taken over the realization of the potential traveler's idiosyncratic taste ω_i. To keep the theoretical part concise, we do not provide a complete characterization of the solution to the dynamic problem. Instead, we illustrate the interplay of prices and ratings as well as the resulting dynamic incentives arising for the strategic host, as these are sufficient to guide our empirical analysis.

Strategic pricing incentives. To illustrate the strategic pricing incentives, consider the first-order condition arising from the Bellman equation:

$$
\begin{aligned}
0 = & \ 1 + \mu(\bar{\Psi}_t) - 2p_t & (7) \\
& + \delta\rho \, (1+\mu(\bar{\Psi}_t)-p_t) \frac{d}{dp_t} \left(E\left[V_{t+1}(\lambda_{t+1}) \middle| \, \omega_i \geq p_t - \mu(\bar{\Psi}_t) \right] \right) & (8) \\
& + \delta\rho \left(V_{t+1}(\lambda_t) - E\left[V_{t+1}(\lambda_{t+1}) \middle| \, \omega_i \geq p_t - \mu(\bar{\Psi}_t) \right] \right) & (9)
\end{aligned}
$$

In contrast to the myopic host, a strategic host not only takes into account the impact of the price on flow profits (7) but also two additional effects.

The first effect relates to the intensive margin effect of prices on rating updates and is captured by (8). A price change has two effects on the expected review scores. First, a higher price directly lowers the expected review via the value-for-money component of the review function—the *value-for-money effect*. Second, a higher price increases the taste of the marginal consumer and, therefore, increases the expected review—the *selection effect*. The value-for-money effect follows directly from our assumptions on the generation of reviews (see Assumption 1): $\frac{\partial}{\partial p} E[\psi_i \mid \theta, \omega, p] < 0$. The selection effect is more subtle. Increasing the price increases the idiosyncratic taste required to induce a booking. Therefore, the average traveler *conditional on booking* has a higher taste for the apartment and the expected review increases as $\frac{\partial}{\partial \omega} E[\psi \mid \theta, \omega, p] > 0$. The overall effect of a price change on the expected review depends on the relative strength of these two effects. Consequently, higher prices are associated with higher reviews whenever (8) is positive. We refer to this effect as the *induced review effect*.

The second relates to the effect of the price on the likelihood of receiving an additional review. Lowering the price increases the probability of a booking and, thus, of receiving an additional review which causes an update in the aggregate rating. If the continuation value, in expectation, increases with an additional rating, the strategic host has an incentive to lower the price to accumulate additional reviews.[9] This extensive margin effect on the rating is captured by (9) and we refer to it as the *rating update effect.*

The following proposition summarizes the strategic pricing incentives and compares the pricing decision of a strategic host with that of a myopic host.[10]

Proposition 1 *The impact of price changes on flow profits determines the pricing incentives of a myopic host. For a strategic host, three considerations determine the pricing incentives: (i) the impact of price changes on flow profits, (ii) the rating update effect, and (iii) the induced review effect.*

(a) The induced review effect *is negative, i.e., puts downward pressure on the price, if and only if the value-for-money effect dominates the selection effect:*

$$\frac{d}{dp_t}\left(E\left[V_{t+1}\left(\lambda_{t+1}\right)\middle|\,\omega_i \geq p_t - \mu(\bar{\Psi}_t)\right]\right) < 0.$$

Otherwise, the induced review effect *is positive and puts upward pressure on the price.*

(b) The rating update effect *is negative, i.e., puts downward pressure on the price, if and only if the expected continuation value following a review is higher than that at the current rating:*

$$V_{t+1}(\lambda_t) < E\left[V_{t+1}\left(\lambda_{t+1}\right) \mid \omega_i \geq p_t - \mu(\bar{\Psi}_t)\right].$$

Otherwise, the rating update effect *is positive and puts upward pressure on the price.*

(c) A strategic host charges a lower price than a myopic host if the value-for-money effect dominates the selection effect (negative induced review effect) and the continuation value is expected to increase at the myopic price (negative rating update effect).

(d) Conversely, if the selection effect dominates the value-for-money effect (positive induced review effect) and the continuation value is expected to decrease at the myopic price (positive rating update effect), a strategic host charges a higher price than a myopic host.

[9] Note that the expected continuation value depends on both the expectation about the incoming reviews and how the continuation profits respond to the respective incoming reviews. Even if a low review is more likely, the expected continuation value might increase because the (less likely) good review leads to a more than proportional increase in profits; e.g., because it leads to a higher displayed rating $\bar{\Psi}$. Conversely, the expected continuation value can decrease even when a good review is more likely.

[10] In the following proposition, we assume that the dynamic pricing problem is sufficiently well-behaved such that the reasoning via the first-order condition with respect to the price is valid.

(e) *If the* rating update effect *and the* induced review effect *go in opposite directions, the overall impact on the price of a strategic host relative to a myopic host is ambiguous.*

Proof. (a) follows immediately from the preceding discussion. For (b), it is straightforward that the sign of (9) depends only on whether the continuation value, $E[V(\lambda_{t+1})|\omega_i \geq p_t - \mu(\bar{\Psi}_t)]$, exceeds the continuation value at the current rating, $V(\lambda_t)$. (c) to (e) are straightforward combinations of the possible cases regarding (a) and (b). ■

2.2 Model Predictions and Numerical Illustration

Before we proceed with the empirical analysis, we briefly summarize the theoretical predictions and corresponding testable hypotheses. Additionally, we present numerical simulations to illustrate these predictions.

Model Predictions. Proposition 1 suggests that strategic pricing behavior hinges on how prices affect ratings. As noted in Proposition 1.(a), this relationship is influenced by the relative strength of the value-for-money effect versus the selection effect. The dominant effect determines whether prices positively or negatively impact induced reviews. In Section 4.1, we evaluate the sign of this relationship. Upon presenting evidence of a negative effect of prices on ratings—which is consistent with Luca and Reshef (2021)—we explore additional hypotheses that derive from this finding.

First, strategic hosts may lower prices to cross salient rating thresholds, thereby enhancing perceived quality by prospective customers. Section 4.2 investigates host pricing around thresholds, finding evidence in line with strategic host behavior. Second, and in line with both strategic behavior and the induced review effect, hosts may benefit from offering discounts when entering the market. The dynamics of market entry, a key focus of our empirical analysis, are examined in Section 5. We show that hosts benefit from offering entry discounts compared to those who do not. Furthermore, evidence indicates that discounts are used more frequently by hosts who appear more professional: this finding is in line with sophisticated hosts using discounts to boost ratings and revenue.

Although the observed correlation between sophistication and discounts aligns with our theoretical framework, it raises the question of whether entry-price dynamics could be due to host-specific factors unrelated to pricing decisions. We address this and related concerns in Section 5.3, where we discuss, among other things, an IV strategy leveraging host-independent variations in discounts.[11]

Numerical Illustration. To illustrate the predictions of our model regarding optimal entry discounts, we implement a parameterized version of the model numerically. The implementation presented here features a dominant value-for-money effect where

[11] As the dynamic effects of price changes are independent of their source, also random discounts should allow insights into the effects of strategic discounts.

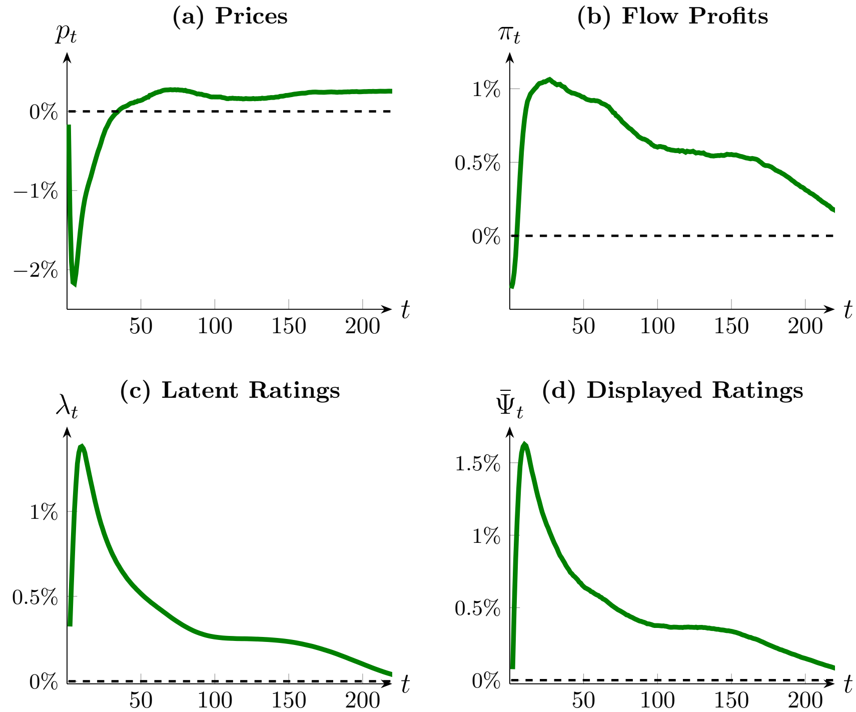

Figure 1: Comparison of strategic and myopic pricing. *The plots indicate the relative difference between the means of different variables over 250,000 simulations for the dynamically optimal pricing strategy and the myopically optimal pricing strategy. The graphs average both the paths for low- and for high-quality hosts. The time horizon for both the computation of the optimal strategy and for the simulations is* $T = 500$ *periods. Additional details of our approach are documented in Appendix A.*

lower prices induce higher realized review scores—therefore, the *induced review effect* incentivizes a price decrease.[12] Figure 6 plots price, flow profit, and rating differences between the dynamically and myopically optimal pricing strategies.

The negative sign of the *induced review effect* determines the overall strategic price-setting in early periods. A strategic host has an incentive to lower the price as to induce higher reviews and thus higher future ratings which imply higher future profits. This price decrease, exemplified by the initially negative price difference in Figure 6a, sacrifices flow profits initially (see Figure 6b). At the same time, it has a positive impact on the rating stock (see Figure 6c and Figure 6d) which quickly allows the firm to reap higher flow profits. However, this effect is only temporary, and over time, the rating and profit differences disappear.

3 The Data

For our empirical analysis, we combine transactions and ratings data on Airbnb listings in Paris, France. Our observations span the entire year 2017.

[12]We outline the detailed implementation in Appendix A.

The transactions data we use for our study are obtained from AirDNA, a specialist for short-term rental data.[13] These data contain information that has been web scraped from the Airbnb platform. They allow us to determine whether a listing was available or booked on a particular date and also to access the corresponding booking price or daily asking price for listings that were available but not booked. Consecutive days of occupancy by the same guests are identified by booking identifiers.

Table 1: Summary Statistics

	N	Mean	Min	Median	Max
Days Available	408,930	23.50	1.00	26.00	38.00
Days Booked	408,930	12.57	0.00	11.00	38.00
Number of Bookings	408,930	2.86	0.00	2.00	30.00
Number of Reviewed Bookings	408,930	1.17	0.00	0.00	22.00
Offer Price	408,930	79.86	1.33	61.83	500.00
Booking Price (All Bookings)	311,393	75.56	1.35	60.00	624.71
Booking Price (Reviewed Bookings)	181,872	71.66	1.19	57.75	1035.00
Overall Rating (Granular Measure)	350,153	92.42	20.00	94.00	100.00
Overall Rating (Stars)	350,153	4.65	1.00	4.5	5.00

Notes: Observations were aggregated for the listing month. For example, the number of available days were calculated as the average number of days a listing was available for booking between two consecutive ratings updates. The time frame between the two consecutive ratings updates corresponds to roughly one month. Depending on the scraping routine, the exact time interval can exceed one month. This explains why the maximum number of available and booked days exceeds 31 days.

The ratings data are obtained from InsideAirbnb.com.[14] These data are web scraped from the Airbnb platform at monthly intervals. At the beginning of each month, updates of the aggregate star ratings for various categories are observed. In total, we observe star ratings in seven different categories, each ranging from one to five stars: (i) overall, (ii) value-for-money, (iii) cleanliness, (iv) check-in, (v) location, (vi) accuracy of the description, and (vii) communication.

The overall rating assesses a traveler's overall experience during a stay in a particular listing and is displayed most prominently to potential guests on the Airbnb website. The other rating categories capture specific aspects of the stay. These are only seen by guests who browse through the accommodation's listing more thoroughly. It should be noted that the overall rating is not a mechanical average of the other rating categories but can be freely chosen by the customer.[15]

While we observe booking information at the daily level, the rating information is only available at the monthly level. Therefore, we cannot directly match transaction prices to corresponding review scores. Instead, we relate the average booking price for a listing in a given month to the aggregate rating at the beginning of the next month (when the rating information in the InsideAirbnb data is provided). Importantly, leaving a review is optional so not all stays are rated. To discard bookings that

[13]See https://www.airdna.co (last accessed: August 16, 2022).

[14]See http://insideairbnb.com (last accessed: August 16, 2022).

[15]The discussion in https://community.withairbnb.com/t5/Hosting/5-stars-in-all-categories-but-4-star-stay/td-p/6934705 (last accessed: August 16, 2022) clarifies that the overall rating is not a mechanical function of the other rating categories.

did not receive a review, we use timestamp information revealing when reviews were submitted.[16] Guests can leave a review on Airbnb within 14 days of their stay. Therefore, if a rating update appears within 14 days of a booking, we label this booking as "reviewed." If there are multiple bookings in the 14 days prior to the review, we choose the closest one to the review date. When studying the impact of prices on ratings, we calculate the monthly average price using only reviewed bookings.

Table 1 shows summary statistics for the main variables of interest. We only include observations for which rating updates are observed in two adjacent months. We otherwise include all listings in our sample. These listings include entire homes as well as private and shared rooms. In our main empirical specification, we focus on how within-listing price variation impacts the rating by including listing fixed-effects. These fixed effects capture the effects of the listing type as well as other time-constant differences on ratings. Changing the sample composition by, e.g., focusing on entire homes, does not alter our qualitative results.

On average, listings are available for 24 days and booked for 13 days per month. There are, on average, 2.86 bookings per month, of which 1.17 are reviewed.[17] We remove observations above the 99th percentile of the price distribution and normalize all prices by the number of guests included in the price.[18] The offer price is the weighted average of the prices observed for the days the listing was booked and the posted prices observed for the days when the listing was available. The average price of reviewed bookings is 71.66 euros per night. The summary statistics for the overall star rating reveal that a majority of the observations enjoy an overall rating of 4.5 or higher. Star ratings (for the overall rating and all other categories) can take on nine values from one to five in half-star steps.

Importantly, the overall star rating is computed based on a more granular rating ranging between 20 and 100, which travelers did not observe when our data was sampled.[19] The number of stars shown to customers is a step function of the underlying granular rating measure: If we denote by r the granular measure, the number of stars, $f(r)$, shown to a potential traveler follows the following rule: $f(r) = 1$ if $r \in [20, 25)$, $f(r) = 1.5$ if $r \in [25, 35)$, $f(r) = 2$ if $r \in [35, 45)$, $\cdots$, $f(r) = 4.5$ if $r \in [85, 95)$, and $f(r) = 5$ if $r \in [95, 100]$.

[16] InsideAirbnb provides auxiliary data about when individual reviews were posted. Thus, we are able to link bookings to subsequent reviews but cannot see the score associated with a given review as this information is not contained in the InsideAirbnb data. We thus need to rely on the aggregate score at the beginning of each month.

[17] A study by Fradkin et al. (2021) reports that 68 percent of bookings receive reviews. However, the geographical scope of their data is not specified. The ratio of reviews to bookings may vary significantly across time and geographies. Additionally, a study on the short term rental market by the Budget and Legislative Analyst's Office of San Francisco notes a discrepancy between the observed ratio of 30 percent and the ratio of 72 percent reported by Airbnb, see page 49 under the following link: https://sfbos.org/sites/default/files/FileCenter/Documents/52601-BLA.ShortTermRentals.051315.pdf (last accessed: August 16, 2022).

[18] The InsideAirbnb data provide information on how many travelers are included in the price a host posts on the Airbnb website.

[19] This has changed since. In 2025, travelers observe a granular numerical rating score.

Airbnb itself does not disclose exactly how individual ratings are aggregated over time. However, the data suggest that the aggregate rating is obtained by simply averaging individual review scores, which is in line with anecdotal evidence.[20] This implies that any effect of prices on ratings should be less pronounced for listings that received more reviews because the marginal impact of one individual review on the average rating diminishes as more ratings accumulate.

4 Induced Review Effect and Strategic Pricing

4.1 Sign of Induced Review Effect

Our model predictions depend crucially on the overall impact of prices on ratings, i.e., the sign of the *induced review effect*. Therefore, we first provide evidence that the value-for-money effect dominates the selection effect for most rating categories.[21] Importantly, we can assess the *induced review effect* separately from the *rating update effect* because it solely depends on the relationship between observed prices—irrespective of how they are set—and induced review scores.

In Figure 2, bars of the same color show the distribution of the overall star rating for listings in the first, second, and third price tercile, respectively. Figure 2 reveals that listings in lower price terciles have systematically lower ratings than listings in higher price terciles. For example, more than 50 percent of the listings in the highest price tercile have the highest possible rating. In the lowest price tercile, only 40 percent of the listings have a five-star rating.

This cross-sectional observation allows for many potential explanations. First, rating differences could be explained by quality differences between apartments and thus be independent of prices when quality is properly accounted for. Second, there could be reverse causality, such that high ratings lead to high prices. Third, the selection effect could dominate the value-for-money effect leading to higher ratings for higher-priced listings.

In the following empirical analysis, we address the first potential explanation by controlling for unobserved time-constant quality differences between listings by including listing-specific fixed effects. We use our matched rating-price pairs to address reverse causality and regress period t aggregate ratings on period $t-1$ prices. We argue that, after taking into account these two potential explanations, the remaining conditional correlation between ratings and prices is indicative of the relative importance of the value-for-money and the selection effect.

[20] See https://airhostsforum.com/t/how-exactly-is-the-star-rating-calculated/14575m (last accessed: August 16, 2022) for a discussion among Airbnb hosts which indicates that the aggregate rating is a simple average of the individual reviews. We provide supporting empirical evidence in Appendix B.

[21] Our theoretical and empirical analysis builds on the assumption that ratings affect demand. Establishing this link between ratings and demand is not our main focus; therefore, in the main text, we do not present results pertaining to the impact of ratings on demand. Previous literature on the subject has established that better ratings positively affect demand (see Section 4 of Tadelis, 2016, for an extensive review). In Appendix C, we present evidence using our data that is consistent with this finding.

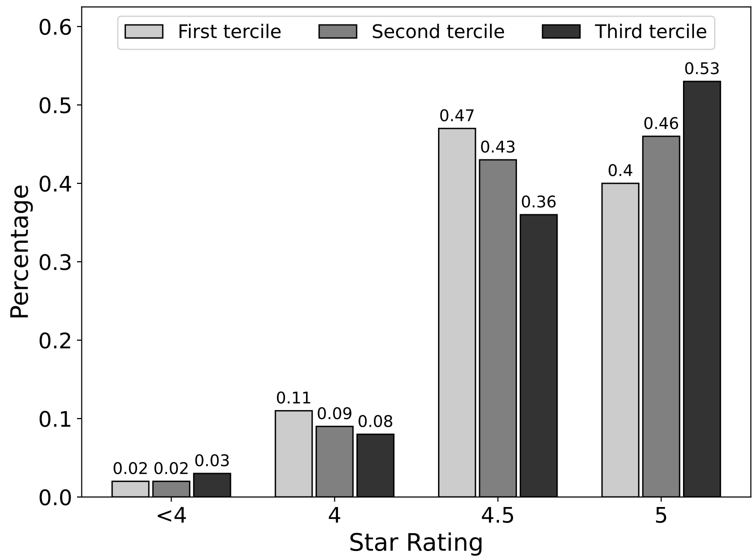

Figure 2: Rating Distribution for Different Price Terciles

Notes: Price terciles are computed based on average listing prices. For exposition, we group ratings below 4 stars into a single category.

Specifically, we estimate the following equation:

$$r_{it}^{cat} = \beta_0 + \beta_1 \times log(p_{it-1}) + X_{it}'\gamma + \mu_i + \mu_t + \epsilon_{it}, \tag{10}$$

where r_{it}^{cat} denotes the aggregate star rating for rating category *cat* of listing i at the start of month t; $log(p_{it-1})$ denotes the logged average price of reviewed bookings in the month prior to observing the aggregate rating; the subscript $t-1$ for the prices emphasizes that we match transaction prices to the ratings we observe immediately after the transactions. This time lag between prices and ratings helps address reverse causality concerns. To empirically assess the sign of the *induced review effect*, our main interest lies in the coefficient β_1. The listing fixed effects μ_i account for time-invariant quality differences across listings, while μ_t denotes the month fixed effects.

X_{it} accounts for time-variant factors. To control for host effort—which can be interpreted as a time-varying quality component—we include information on the host's response rate (i.e., how often does the host respond to inquiries from potential guests), which is expressed as a share between zero and one. To account for the averaging in the calculation of the aggregate ratings, we also include the number of ratings and its square in X_{it}.

The top panel of Table 2 shows the main coefficients of interest from estimating Equation (10) for each of the seven rating categories. The results appear intuitive in light of our model. The price coefficient is negative and the largest in absolute terms for the value-for-money rating. All other rating categories appear less affected by prices. Interestingly, the location rating, which asks for an arguably time-invariant quality aspect of the listing, is among the least correlated with prices. Crucially, the price negatively correlates with the overall rating, which is the most salient rating

Table 2: Price-Rating Regressions: Value-for-money vs. selection effect.

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
All Listings							
log(price)	-0.026^{***}	-0.035^{***}	0.003	-0.025^{**}	-0.003	-0.018^{**}	-0.009
	(0.006)	(0.008)	(0.006)	(0.008)	(0.009)	(0.007)	(0.007)
R^2 Adj.	0.890	0.805	0.841	0.814	0.862	0.804	0.806
R^2 Within Adj.	0.002	0.001	0.000	0.000	0.001	0.001	0.000
Split by Price Level							
log(price) - low	-0.044^{***}	-0.058^{***}	0.013	-0.044^{**}	-0.006	-0.012	-0.006
	(0.012)	(0.015)	(0.012)	(0.016)	(0.016)	(0.013)	(0.012)
log(price) - high	-0.017^{*}	-0.025^{**}.	-0.002	-0.016^{*}	-0.001	-0.020^{*}	-0.011
	(0.007)	(0.009)	(0.007)	(0.008)	(0.010)	(0.008)	(0.008)
R^2 Adj.	0.890	0.805	0.841	0.814	0.862	0.804	0.806
R^2 Within Adj.	0.002	0.001	0.000	0.000	0.001	0.001	0.000
Obs.	179,972	179,734	179,749	179,901	179,962	179,863	179,773

Notes: $*$, $**$, $***$ indicate statistical significance at the five, one, and 0.1 percent levels, respectively. Standard errors are shown in parentheses and are clustered at the listing level. All regressions control for listing fixed effects, month fixed effects, a second-order polynomial of the number of ratings, and the host response rate as a proxy for the effort of the host. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively.

to customers. This suggests that the value-for-money effect indeed dominates the selection effect in our data.[22] We conjecture that the accuracy rating is particularly sensitive to the difference between the expected and the actual quality of a listing. Because of the importance of the quality dimension, the strong negative correlation between prices and the accuracy rating appears intuitive in the presence of a dominant value for money effect.

The bottom panel of Table 2 extends the analysis by interacting the price coefficients with a dummy variable that splits listings into a below- and above-median-price category. This categorization is based on each listing's average price, so the dummy variable is constant for every listing. The results in Table 2 are consistent with a higher price sensitivity in the ratings of price-conscious customers who book less expensive listings. The increased price sensitivity, in turn, points to a stronger value-for-money effect—and thus a more negative induced review effect—of prices on ratings.[23]

We conclude this section by noting that our results do not preclude the presence of a selection effect in the determination of the sign of the *induced review effect*. However, our results suggest that the value-for-money effect dominates the selection

[22] There could be additional within-listing unobserved variation in quality over time that affects prices as well as ratings. For example, a host might renovate an apartment, resulting in higher quality but also higher prices. We would expect this type of bias to result in a higher likelihood of finding a positive relationship between prices and ratings, making it harder to detect a dominant value-for-money effect.

[23] We present robustness checks based on a first-differences estimator in Appendix D.

effect and leads to an overall incentive to decrease prices as this on average increases ratings. Based on our theoretical framework, these insights allow us to move on to investigate dynamic price and ratings patterns, as they determine the hosts' incentives and thus the predicted sign of the relationships between the variables of interest.

4.2 Strategic Pricing Behavior

According to the theory, the *negative induced review effect* identified in Section 4.1 implies that sophisticated hosts can improve their ratings by reducing prices. We now provide empirical evidence in line with strategic hosts seeking to improve and maintain higher ratings through lower prices.

The motivation to lower prices is based on weighing the immediate loss in revenue due to lower prices against the future revenue gains from increased demand. During the sample period, Airbnb ratings were displayed to potential travelers in discrete increments of half-stars, ranging from 1 to 5. However, the actual average rating is a continuous value that hosts can calculate. Intuitively, hosts benefit substantially when consumers see a rating of 4.5 instead of 4 stars and thus have strong incentives to cross the threshold that triggers the 4.5-star display. This creates an incentive to forego revenues in the present period by lowering prices to receive better ratings, jump the salient threshold, and benefit in the future.[24] The closer to the salient threshold, the stronger the incentive to decrease prices is for strategic hosts because they are more likely to be able to benefit in the nearer future. A numerical implementation of our model in Appendix A.1 confirms that this behavior is indeed what we would expect of strategic hosts in the model.

Moreover, the *rating update effect* is likely to provide an additional incentive to lower prices when listings are just below a threshold—this is because a positive review realization pushing the listing beyond the threshold has a much higher impact on the continuation value than a negative review realization keeping it below. This argument is in line with the empirical analyses by Sorokin (2021) in the context of video games on Steam, who finds that sellers close to moving up salient review tiers are more likely to discount, and by Zhong (2022).[25]

To analyze whether our data support this hypothesis, we focus on the subset of active listings whose hosts are more likely to engage in strategic pricing behavior.[26] We estimate the following equation:

$$p_{it} = \beta_0 + \sum_{j=lb}^{ub} \beta_j \mathcal{I}(r_{it} = j) + q_i + l_i + \mu_t + \epsilon_{it}\,. \tag{11}$$

p_{it} denotes the offer price of listing i in period t, r_{it} denotes the granular overall rating

[24] We provide evidence for the revenue impact of crossing salient rating thresholds in Appendix C.

[25] Because of the rating update effect, similar price patterns are possible even with a dominant selection effect, see Appendix A.3. However, the price decreases just before the threshold are muted, while price increases once the threshold is passed are exacerbated.

[26] We include a listing if the number of annual bookings is larger than the median observed across listings. We expect less active, occasional hosts to be less likely to engage in strategic pricing.

of listing i at the *beginning* of period t,[27] q_i denotes a five-scale quality indicator variable, l_i denotes location-specific fixed effects, μ_t denotes month fixed effects, and β_j captures the conditional average price of a listing with a granular rating $j \in [lb, ub]$ compared to the baseline granular rating category β_0. The quality indicators q_i are obtained from a hedonic regression of offer prices on listing fixed effects and month fixed effects and correspond to the quintiles of the distribution of the listing fixed effects.[28]

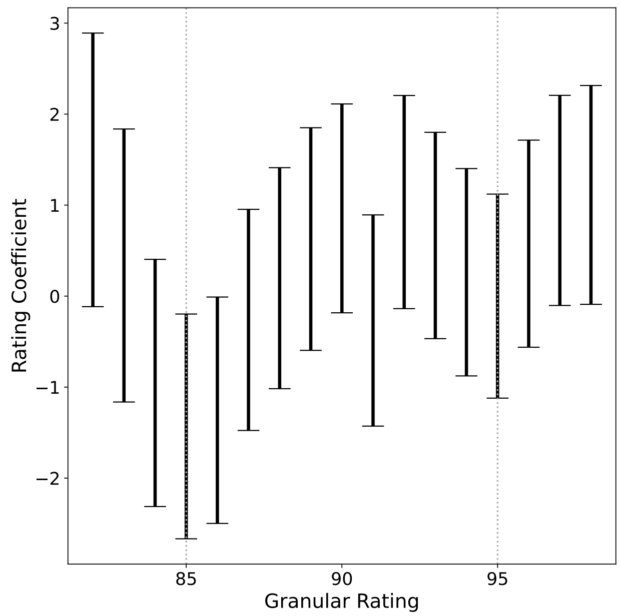

Figure 3: Conditional Average Offer Prices Around Salient Threshold

Notes: The 95% confidence intervals are based on heteroskedasticity-robust standard errors.

Figure 3 shows the rating coefficients we obtain when estimating Equation (11) for $j \in [82, 98]$ (using the granular rating of 81 as the baseline category).[29] The range has been selected to cover all ratings in the vicinity of the four- to four-and-half, and four-and-a-half to five-star thresholds (85 and 95, respectively).

Figure 3 reveals a V-shaped pricing pattern around the salient ratings thresholds. The monotonic price decreases observed before each threshold are consistent with the mechanism outlined above. Similarly, the price increases after each threshold are consistent with hosts seeking to exploit the benefits of higher ratings. Note that prices appear to increase only gradually after the thresholds. This suggests that hosts are aware that raising prices too quickly after crossing a threshold might lead them to

[27]Note that, so far, we focus on prices associated with bookings. In this analysis, we instead focus on offered prices, i.e., the weighted average of the prices for the days a listing was booked and the days a listing was available. The reason for this change is that we are interested in host behavior in this analysis, whereas the previous analyses focused on guest behavior.

[28]In Appendix E, we provide more details on the estimation of the quality fixed effects and discuss the issues hampering an analysis using conventional listing-specific fixed effects.

[29]Table 8 in Appendix E shows the corresponding estimates.

lose their newly obtained higher salient ratings due to the dominant value-for-money effect.[30]

5 Price-Rating Dynamics

In this section, we turn to an analysis of the dynamics in the relationship between prices and ratings. All else equal, the dominant value-for-money effect and the evidence for strategic pricing behavior established by our previous analyses implies that strategic hosts should offer relative discounts when entering the platform. This allows them to generate better ratings than hosts charging higher entry prices. As a result, these hosts should be able to charge higher prices once their ratings have stabilized. Recalling Figure 6, this reasoning is supported by our numerical implementation. Sacrificing flow profits upon entry induces higher rating stocks and, thus, higher flow profits in the medium run.

To empirically evaluate these dynamics, in Section 5.1, we start by providing evidence that listings offering relatively lower prices upon entering the platform benefit from this "entry price discount" in subsequent periods. Consequently, we focus on the effects of entry discounts rather than the specific origin of entry price variations. In Section 5.2, we present evidence suggesting that more professional hosts employ such entry discounts, supporting the concept of strategic pricing reflected in our theoretical model predictions. Finally, in Section 5.3, we address endogeneity concerns related to the differences between hosts who offer discounts and those who do not.

5.1 The impact of entry price discounts

To analyze the impact of entry price discounts, we restrict the sample to new entrants and only use listings that do not have more than six ratings when we first observe them.[31] We estimate the following equation:

$$y_{it} = \sum_{\tau=1}^{6}(\alpha_\tau + \beta_\tau \times d_i)\mathcal{I}(m(i,t) = \tau) + X_{it}'\gamma + \epsilon_{it}\,, \tag{12}$$

where y_{it} denotes different outcome variables for listing i at time t. The outcome variables include the price, the number of bookings, the different rating categories, and the revenues. d_i measures the initial discount listing i offers in the first period it enters the market. $\mathcal{I}(m(i,t) = \tau)$ is a dummy variable that takes the value one if month t marks the τth period since the entry of listing i. The idea of interacting the cross-sectional variable, d_i, with period fixed effects, $\mathcal{I}(m(i,t) = \tau)$, is inspired by Huber et al. (2021). The coefficients β_τ show cross-sectional differences between listings as a function of the size of the entry price discount τ periods after entry.

[30] We provide additional robustness checks for the results shown in Figure 3 in Appendix E.

[31] To have a sufficient number of observations to study changes over time, we restrict attention to listings that we observe for at least six months, with bookings in each of these six months. To have a balanced panel, we only include the first six monthly observations for each listing.

X_{it} subsumes time-varying control variables and fixed effects. We always control for month fixed effects, location fixed effects, and a fixed effect for the month-of-market-entry.[32] Additionally, when the outcome variable is the number of bookings or total monthly revenues, the total number of days a listing was available and the number of reviews at the beginning of the respective month are also included in X_{it}. Including the number of reviews allows us to control for the potential demand effect of increasing the number of reviews, as discussed by Vellodi (2024). Note that the inclusion of listing fixed effects would not allow us to study the relationship between y_{it} and the initial discount in every period. Instead, it would only allow us to study how the relationship between y_{it} and the initial discount *changes* across periods. In Appendix F, we show that the results obtained when including listing fixed effects in Equation (13) are consistent with the findings presented here.

To measure the initial discount, we compare the price charged in the first month to the average price charged in subsequent months. Formally, we calculate

$$d_i = 1 - \frac{p_{i1}}{\bar{p}_{i2-6}}\,, \tag{13}$$

where $\bar{p}_{i2-6}$ denotes the average price charged after the first month. A negative value of d_i arises if the initial price charged is higher than the average price charged in subsequent periods. A positive value captures an initial discount. Note that the maximum discount is naturally one. Accordingly, β_τ in Equation (12) can be interpreted as the effect of a 100 percent first-period discount on the outcome τ periods after entry. To aid the interpretation of the following results, we report the estimates scaled by the median discount of seven percent.[33]

Figure 4 shows the estimated β_τ-coefficients multiplied by the median discount for the different outcome variables.[34] According to Figure 4a, listings with a median entry discount of seven percent first charge a five euro lower price and then a three euro higher price compared to listings without entry discount. In this context, it is important to emphasize that this reversal is not pre-determined, as nothing prevents any of the curves in Figure 4 from lying strictly below or above zero throughout. For example, a curve lying above zero throughout in Figure 4a and Figure 4c would mean that listings with discounts are systematically more expensive and simultaneously more booked which, in turn, should immediately raise serious concerns related to a systematically higher unobserved quality of listings with discounts.

While this type of endogeneity cannot be fully ruled out, the sign reversal for prices and the pattern observed for bookings assuage concerns related to unobserved quality differentials that would allow listings with discount to charge higher prices without incurring a demand-side penalty from it in the longer run. This insight is further corroborated by the fact that we observe this downward sloping pattern for the value-for-money and the overall rating but not for rating categories that are less

[32]The location fixed effects are based on statistical, geographical units called IRIS, which were defined by the French National Statistical Office (INSEE) to capture areas with similar population sizes. We have approximately 990 different IRISs in our sample.

[33]The median discount is calculated conditional on a listing setting a discount larger than zero.

[34]We provide the detailed regression results in table format in Appendix F

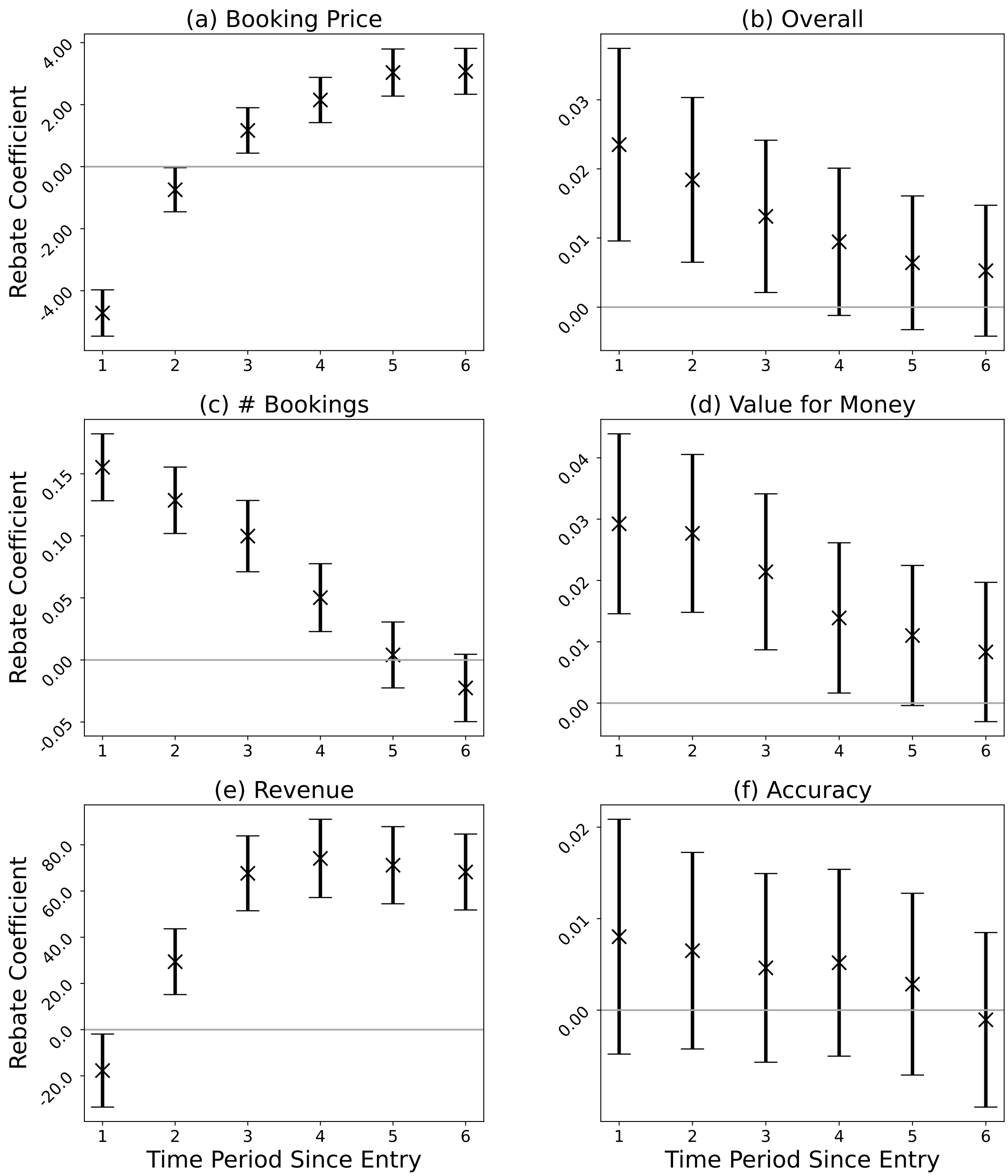

Figure 4: β_τ for Different Dependent Variables (Scaled at Median Discount of 7%)

Notes: The 95% confidence intervals are based on standard errors clustered at the listing level.

likely associated with the price. In Figure 4f, we report results for the accuracy rating, because it is likely particularly sensitive to quality-related aspects (as discussed in Section 4.1). We provide the results for other rating categories in Appendix F.

Figure 4b shows that an initial discount is correlated with a higher overall rating. In the first months, listings that offer an initial discount receive, on average, better overall ratings. This is consistent with the dominant value-for-money effect and further corroborated by Figure 4d, which shows a similar pattern for the value-for-money rating. When the prices increase in subsequent periods, this advantage vanishes. Importantly, even though hosts who offer an initial discount charge higher prices, there is no penalty tied to the value-for-money rating after six months compared to hosts who do not offer an entry discount. This is consistent with the initial discount offering an initial boost in the rating score, which dissipates only gradually.

Consistent with a downward sloping demand curve, Figure 4c shows that the initial discount seems to draw additional bookings in the first periods. However, even though the prices of listings with initial discounts are relatively higher in later periods, they do not experience lower numbers of bookings. Finally, as a combination of the results for the transaction prices and the number of bookings would suggest, listings that set an initial discount can generate higher revenues in subsequent periods (see Figure 4e). At a median initial discount, this increase in medium-run revenues amounts to approximately 68 euros, or three percent, each month. We emphasize that we would expect these differences to dissipate in the long run, as is predicted by the theory and corroborated by the trends in Figure 4c and Figure 4d.

5.2 Who sets entry discounts?

We anticipate that strategic hosts are more likely to use entry discounts when entering the platform. While there is no clear measure of hosts' strategic behavior, it is likely that more professional hosts are also more likely to behave strategically. Therefore, we examine how various indicators of host professionalism correlate with their discount-setting behavior.

We focus on the entry sample used in Figure 4. For each listing, we use the observation at the time of entry, making this a cross-sectional analysis. We perform a regression analysis with a dummy variable set to one if the listing offers an entry discount, using measures of host professionalism as the independent variables. These measures include whether a listing is instant-bookable, whether the host manages multiple listings, whether the listing involves a separate cleaning fee, whether a security deposit is required, and whether the host is designated as a "superhost".

Table 3 presents the results of this analysis. Columns (1) through (6) analyze each professionalism measure separately, while the last column combines all variables. The coefficients indicate the percentage point increase in the likelihood of observing an entry price discount. The findings suggest that more professional hosts are more likely to offer an entry discount. Specifically, instant-bookable listings are 8 percentage points more likely, listings managed by multi-hosts are 4 percentage points more likely, those with cleaning fees are 5 percentage points more likely, those requiring a deposit are 2 percentage points more likely, and superhost listings are 4 percentage points more likely to have an entry discount. The effect size of each indicator diminishes in the last specification due to collinearity between variables.

Table 3: Regression of "has discount" dummy on proxies for professionalism

	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.534***	0.548***	0.523***	0.545***	0.559***	0.494***
	(0.007)	(0.007)	(0.010)	(0.009)	(0.006)	(0.012)
Inst. book.	0.083***					0.078***
	(0.012)					(0.012)
Multi-host		0.036**				0.016
		(0.012)				(0.012)
Clean. fee			0.052**			0.047***
			(0.012)			(0.013)
Deposit				0.024*		0.004
				(0.011)		(0.013)
Superhost					0.043	0.023
					(0.033)	(0.034)
Num.Obs.	8,126	8,126	8,126	8,126	8,126	8,126
R2 Adj.	0.006	0.001	0.002	0.000	0.000	0.008

Notes: $*$, $**$, $***$ indicate statistical significance at the five, one, and 0.1 percent levels, respectively. Heteroskedasticity-robust standard errors are shown in parentheses.

5.3 Endogeneity concerns related to host-heterogeneity

Our findings suggest that strategic hosts employ entry-price discounts as part of a forward-looking, profitable strategy. While this aligns with our theoretical framework, it also raises potential endogeneity concerns. Specifically, strategic hosts may use non-price variables differently than non-strategic hosts in a way that impacts the observed outcomes we use in our empirical analysis.

No changes in effort related rating variables over time There is a concern that the results in Figure 4, especially after the initial period, might be influenced by differential changes in non-price attributes. For instance, strategic hosts might systematically exert more effort in later periods, allowing them to maintain good ratings despite higher prices. However, it is important to note that all rating categories other than the overall and value-for-money rating, including those likely capturing host effort, like the communication and cleanliness ratings, appear stable over time (see Appendix F). If effort levels were changing, we would expect to see similar changes in effort-related ratings as observed in the overall and value-for-money ratings. Instead, we see almost no change in the effort-related rating categories, indicating stability over time.

IV-Strategy To address endogeneity concerns related to host heterogeneity, we implement an instrumental variable (IV) strategy using demand proxies to explain the discount in the entry period. Local demand conditions determine the prices that hosts can charge at any point in time. If a listing enters in a period of relatively low demand and demand increases subsequently, this can result in an entry discount driven entirely by variation in local demand. If the demand conditions at the time of entry are independent of host-specific characteristics and ratings, this approach helps separate the initial discount from unobserved host attributes. This assumption is reasonable if it is challenging for hosts to control the exact timing of entry. For example, factors such as the complexity of timing real estate acquisitions and the subsequent refurbishment requirements support this hypothesis, particularly for professional hosts. The relevance of these instruments arises from the general correlation between prices and demand factors. Our detailed approach and results are described in Appendix G. We find no evidence that the patterns shown for the ratings in Figure 4 are altered by the IV approach, which reinforces our confidence that the observed pattern—low entry prices leading to higher ratings and higher profits in the future—represents a causal mechanism consistent with our theoretical framework.

5.4 Quality-Learning Hypothesis

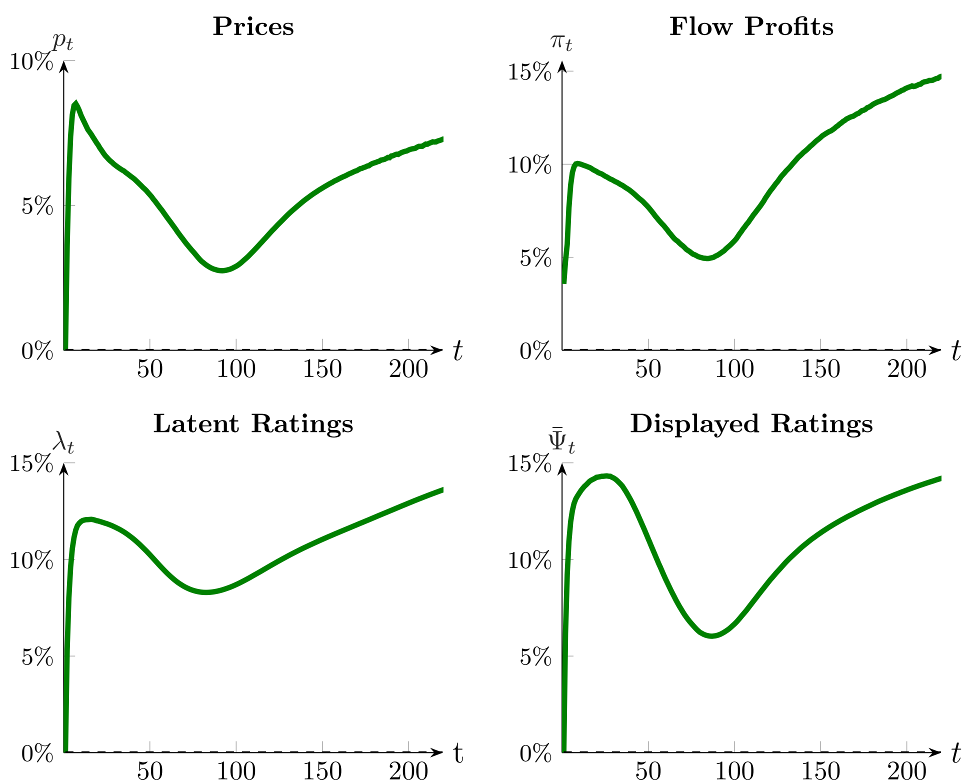

Figure 5: Quality-learning hypothesis: high vs. low quality

The above results on price-rating dynamics are consistent with the hypothesis that hosts who charge an entry discount benefit from an immediate rating advantage, and obtain higher revenues in the medium run. While the patterns observed in Figure 4 correspond well to our model, some of them are potentially also consistent with a competing *quality-learning* explanation: initially unaware hosts might learn about the quality of their own listings after entering the market and, subsequently, adjust their prices.

Under the *quality-learning* hypothesis, hosts with above-average-quality listings initially charge prices which are too low and, after learning their actual quality, increase prices to match their quality. Thus, our empirical specification would classify these as listings offering an initial discount. Analogously, below-average-quality listings would charge high prices upon entry and adjust them downwards in subsequent periods. Such behavior would result in similar patterns as those observed for prices and quantities in Figures 4a and 4c. We illustrate these patterns associated with the quality-learning hypothesis in a numerical implementation of a quality-learning variant of our model (see Appendix A.2 for details).

Figure 5 shows the ratios of simulated price, rating and profit paths, comparing high-quality and low-quality listings. The quality-learning model would indeed exhibit the feature that listings that offer a discount receive better ratings as they have an inherently higher quality. However, under the quality learning hypothesis, we would expect a divergence in the long-run ratings between listings that set an entry discount and listings that do not—this would be driven by the underlying differences in quality. However, these persistent long-run differences, depicted in the two bottom panels of Figure 5, are inconsistent with our empirical results.[35]

Specifically, we do not find strong supporting evidence for long-run discrepancies in the ratings for listings that set discounts (high-quality listings under quality learning) and those who do not (low-quality listings). Therefore, we conclude that the quality-learning hypothesis is not the main force behind the empirical patterns we document.

6 Implications

In this section, we discuss some implications of our analysis for hosts deciding on their pricing strategies and platform operators designing their rating systems.

Implications for Hosts. Our analysis has several implications for hosts on Airbnb. First, the dominant value-for-money effect provides a clear pathway for hosts to improve their ratings by lowering prices. This particularly holds true for listings catering to more price-conscious "budget travelers", where the negative impact of prices on reviews is even more pronounced. Similarly, discounts are more valuable around thresholds such that the salient aggregate rating observed by consumers changes: lower prices increase the likelihood of crossing the threshold following a booking and rating update.

Second, discounts are particularly valuable for new listings, because the simple averaging that Airbnb uses to compute aggregate ratings implies a higher sensitivity of the rating stock to incoming reviews when there are few reviews. This is in

[35] The underlying quality differences are also the driver of the persistent price ratios and flow profit ratios depicted in the upper two panels.

addition to other considerations such as the resolution of the cold start problem where consumers may be less willing to stay at an apartment that has few or no prior reported user experiences. Our empirical analysis shows that entry discounts are worthwhile for hosts of new listings. Relatively lower entry prices lead to better ratings and more bookings in the initial months, which allows higher prices and the realization of higher revenues subsequently.

Implications for Platform Operators. From the perspective of platform operators, strategic seller pricing to influence ratings may be undesirable. For example, it may impede the overall informativeness of the rating system. Our analysis reveals several avenues through which these strategic pricing incentives may be reduced. Refining the sub-categories and the prominence with which they are displayed to users—most notably the value-for-money rating—may help to isolate the impact of prices on reviews and mitigate price distortions due to reputation management.

Additionally, coarse aggregate ratings and the resulting thresholds around which the salient rating observed by consumers changes amplify strategic pricing incentives around thresholds due to their substantial impact on consumers. Notably, Airbnb itself has changed the ratings displayed to consumers in their search results from the coarse aggregate star rating to a granular numerical score. In light of our model and empirical analysis, this reduces strategic pricing incentives for hosts by avoiding discontinuous jumps in perceived apartment quality around salient thresholds.

7 Conclusion

We investigate whether hosts on the short-term accommodation platform Airbnb can influence their ratings through strategic pricing. In a stylized theoretical framework, we illustrate two opposing effects of a higher price on realized review scores. First, higher prices directly result in lower scores due to a reduced value for money. Second, higher prices indirectly result in higher scores due to the self-selection of travelers into booking: only travelers with a strong idiosyncratic preference for the listing will book it. The net effect of prices on realized review scores depends on which of these effects, the value-for-money or the selection effect, dominates.

Using data on Airbnb transactions and corresponding ratings in Paris for 2017, we find that higher prices reduce most rating categories, suggesting that the value-for-money effect dominates the selection effect. The relationship is most prominent for the value-for-money rating. However, we do not find such a relationship for the location rating. We argue that this result is in line with a stronger selection effect for the location rating, which derives from travelers' idiosyncratic preferences for specific locations in the city. Our results suggest that hosts can strategically reduce prices to improve their future ratings.

Given the dominant value-for-money effect, we posit that hosts can benefit from offering discounts upon entry. Our analysis of entry pricing confirms this hypothesis. Listings with relatively lower entry prices receive better ratings and more bookings early on, allowing hosts to charge higher prices and realize higher revenues in subsequent periods. Overall, a median entry discount of seven percent leads to three percent higher revenues six months down the line. We provide evidence that our price variation, at least to some extent, derives from strategic host behavior. Specifically, we show that measures of hosts' professionalism are positively correlated with entry discounts. Moreover, around thresholds such that the salient aggregate rating visible to potential travelers improves discontinuously, hosts appear to strategically lower their prices as to increase the likelihood of crossing the threshold.

It is important to note that hosts' strategic use of prices to influence their ratings may impede the informativeness of rating systems. In our theoretical model, ratings enter the travelers' expectations in a reduced form. In a more elaborate model, travelers could factor in the strategic pricing, which affects how their beliefs react to ratings—potentially lowering the informativeness of the rating system. If this is undesirable for platforms, then providing additional guidance to travelers in the ratings process—for example, mentioning that the price should be ignored in particular rating categories—might restore the system's informativeness. Such insights have important ramifications for the design of online reputation and feedback systems.

References

Bondi, T. (2025). Alone, Together: A Model of Social (Mis)Learning from Consumer Reviews. *Marketing Science (accepted).*

Cabral, L. and Li, L. (2015). A Dollar for Your Thoughts: Feedback-Conditional Rebates on eBay. *Management Science*, 61(9):2052–2063.

Carnehl, C., Stenzel, A., and Schmidt, P. (2024). Pricing for the Stars: Dynamic Pricing in the Presence of Rating Systems. *Management Science*, 70(3).

Culotta, A., Jin, G. Z., Sun, Y., and Wagman, L. (2023). Safety Reviews on Airbnb: An Information Tale. *NBER Working Paper*, 31855.

Fang, L. (2022). The Effects of Online Review Platforms on Restaurant Revenue, Consumer Learning, and Welfare. *Management Science*, 68(11):8116–8143.

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

Foroughifar, M. and Mehta, N. (2023). The Challenges of Deploying an Algorithmic Pricing Tool: Evidence from Airbnb. *Working Paper.*

Fradkin, A., Grewal, E., and Holtz, D. (2021). Reciprocity and Unveiling in Two-Sided Reputation Systems: Evidence from an Experiment on Airbnb. *Marketing Science*, 40:1013–1029.

He, S., Hollenbeck, B., and Proserpio, D. (2022). The Market for Fake Reviews. *Marketing Science*, 41(5):896–921.

Huang, Y. (2022). Pricing Frictions and Platform Remedies: The Case of Airbnb. *Working Paper.*

Huber, K., Lindenthal, V., and Waldinger, F. (2021). Discrimination, Managers, and Firm Performance: Evidence from "Aryanizations" in Nazi Germany. *Journal of Political Economy*, 129(9).

Jeziorski, P. and Michelidaki, K. (2023). How to "buy" honest reviews. Experimental evidence of the impact of prices on online reviews. *Working Paper.*

Jia, J., Jin, G. Z., and Wagman, L. (2021). Platform as a Rule Maker: Evidence from Airbnb's Cancellation Policies. *NBER Working Paper*, 28878.

Johnen, J. and Ng, R. (2024). Harvesting Ratings. *LIDAM Discussion Paper CORE*, 2023/06.

Livingston, J. A. (2005). How Valuable Is a Good Reputation? A Sample Selection Model of Internet Auctions. *Review of Economics and Statistics*, 87(3):453–465.

Luca, M. (2016). Reviews, Reputation, and Revenue: The Case of Yelp.com. *Harvard Business School NOM Unit Working Paper*, 12-016.

Luca, M. and Reshef, O. (2021). The Effect of Price on Firm Reputation. *Management Science*, 67(7):4408–4419.

Luca, M. and Zervas, G. (2016). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. *Management Science*, 62(12):3412–3427.

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

Proserpio, D. and Zervas, G. (2017). Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews. *Marketing Science*, 36(5):645–665.

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

Rossi, M. and Schleef, F. (2024). Quality Disclosures and Disappointment: Evidence from the Academy Nominations. *CESifo Working Paper*, 11573.

Shampanier, K., Mazar, N., and Ariely, D. (2007). Zero as a Special Price: The True Value of Free Products. *Marketing Science*, 26(6):742–757.

Sorokin, D. (2021). Reaching for The Stars: Discounts and Review Tier Transitions in the Video Games Market. *Working Paper*.

Tadelis, S. (2016). Reputation and Feedback Systems in Online Platform Markets. *Annual Review of Economics*, 8:321–340.

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

Vellodi, N. (2024). Ratings Design and Barriers to Entry. *Working Paper*.

Zegners, D. (2019). Building an Online Reputation with Free Content: Evidence from the E-book Market. *Working Paper*.

Zhong, Z. (2022). Chasing Diamonds and Crowns: Consumer Limited Attention and Seller Response. *Management Science*, 68(6):4380–4397.

Appendix

A Numerical Implementation

In the following, we describe the numerical implementation of our model. For the numerical implementation, we focus on a finite horizon and need to parameterize the functional form of the consumer's utility function, the review function, and an updating procedure. In particular, we suppose that the consumer's utility is linear in quality, with a high-quality apartment providing a gross utility of one and a low-quality apartment providing a gross utility of zero. Hence, we obtain

$$U(\bar{\Psi}, \omega_i, p) = \mu_t(\bar{\Psi}_t) + \omega_i - p.$$

The review function is implemented as follows. The probability that a consumer leaves a review of ψ stars depends on her realized utility combined with a price-internalization parameter κ. In particular, we introduce five brackets of realized utilities, $\{[u_0, u_1), [u_1, u_2), [u_2, u_3), [u_3, u_4), [u_4, u_5]\}$. Denote by U_i the interval $[u_{i-1}, u_i)$. If $U(\theta, \omega_i, p) = \theta + \omega_i - \kappa p \in U_i$, the consumer leaves review ψ with probability $p(\psi, U_i)$.

We assume that hosts are aware of the individual reviews that they have received and can compute the average latent rating λ_t. As in our application, we assume that consumers only observe the rating rounded to half-stars, $\bar{\Psi}_t(\lambda_t)$. We normalize reviews and ratings in line with the Airbnb framework during our time period such that a five-star (four-star, etc.) rating contributes a granular score of 100 (80, etc.) to the latent rating average λ_t and compute the thresholds accordingly.

The updating rule of consumers is such that they ignore calendar time and update exclusively based on the observed rating. Consumers believe that high-quality apartments are being rated according to the review function $p(\psi, U_h)$ and low-quality apartments according to the review function $p(\psi, U_l)$. Based on these assumptions, consumers apply a simplified Bayes' rule together with the prior μ_0 to update their beliefs.[36] Note that the observed rating is finer (full and half-stars) than the review function (only full stars). Hence, we obtain for full-star ratings

$$\mu(\bar{\Psi}_t = \psi) = \frac{\mu_0 p(\psi, U_h)}{\mu_0 p(\psi, U_h) + (1 - \mu_0) p(\psi, U_l)}$$

and for low-star ratings

$$\mu(\bar{\Psi}_t = \psi + 0.5) = \frac{\mu_0 (p(\psi, U_h) + p(\psi + 1, U_h))/2}{\mu_0 (p(\psi, U_h) + p(\psi + 1, U_h))/2 + (1 - \mu_0)(p(\psi, U_l) + p(\psi + 1, U_l))/2}.$$

Consumers purchase as in our model only if their expected payoff from booking exceeds the outside value of zero. Conditional on purchase, consumers leave a review with probability ρ according to the review function specified above, where the realizations are drawn independently over time.

[36] In particular, consumers treat the observed rating as a single review, assuming that half star ratings are generated by a mixture of full-star ratings.

We solve for the optimal pricing strategy of the host for each quality realization by backward induction through time with the state variables being the number of reviews received and the latent rating λ_t. In Table 4, we specify the parameters used in the implementation, and in Table 5, we specify the review function used.

Parameter	**Value**
T	500
δ	0.95
μ_0	0.75
ρ	0.41
κ	1
u_0	0
u_1	0.25
u_2	0.3
u_3	0.35
u_4	0.4
u_5	1
U_h	U_5
U_l	U_3

Table 4: Table of Parameters

	$\psi = 1$	$\psi = 2$	$\psi = 3$	$\psi = 4$	$\psi = 5$
$p(\psi, U_1)$	0.15	0.15	0.15	0.15	0.40
$p(\psi, U_2)$	0.075	0.10	0.125	0.25	0.45
$p(\psi, U_3)$	0.04	0.05	0.06	0.30	0.55
$p(\psi, U_4)$	0.02	0.04	0.04	0.30	0.60
$p(\psi, U_5)$	0.002	0.005	0.008	0.185	0.80

Table 5: Review probabilities for utility brackets.

After solving for the optimal strategy, we simulate the model 250,000 times for each quality realization and plot the averages over these simulations and over the qualities based on the prior μ_0. The results are displayed in Figure 6.

To obtain Figure 7, we proceed as follows. For each latent rating λ_t, we compute the average time period and average number of reviews for which they occur in the simulations. Then, we compute the optimal price that corresponds to these average state variables and also obtain the corresponding average price for the given displayed rating. We plot the deviation of the price corresponding to the latent rating relative to the average price for all latent ratings leading to the same displayed rating. In computing the average number of reviews and the average time period, we adjust the state variables within each rating bucket to be close to the average within that bucket, as for some ratings there is substantial variation in the raw averages. In particular, we use a convex combination between individual averages and bucket

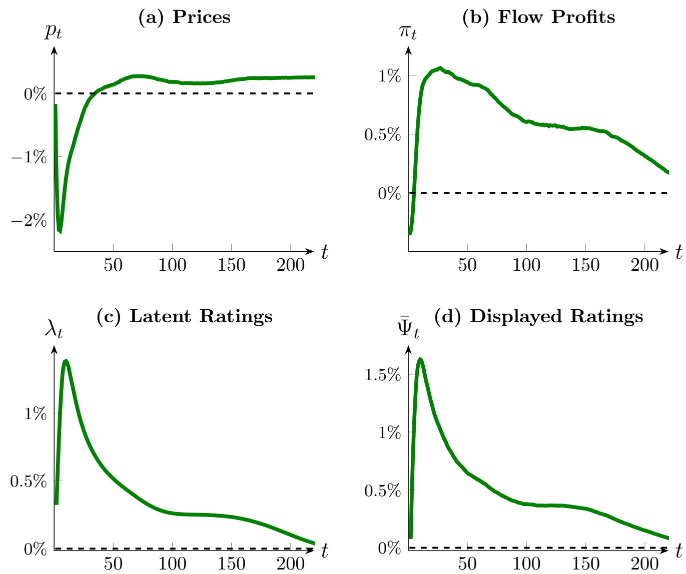

Figure 6: Comparison of strategic and myopic pricing. *The plots indicate the relative difference between the means of different variables over 250,000 simulations for the dynamically optimal pricing strategy and the myopically optimal pricing strategy. The graphs average both the paths for low- and for high-quality hosts. The time horizon for both the computation of the optimal strategy and for the simulations is* $T = 500$ *periods. Additional details of our approach are documented in Appendix A.*

averages with weights 0.1 and 0.9. This ensures that we capture sufficient variation across latent ratings in a bucket but still obtain meaningful comparisons within the rating bucket.

A.1 Pricing at Ratings Thresholds

In this section, we illustrate host incentives when hosts approach salient ratings thresholds. Whenever the latent rating is close to a threshold (in this graph at latent ratings of 85 and 95) such that there is a shift in the salient star rating, strategic hosts have an incentive to lower the price to increase the likelihood of obtaining a sufficiently positive review which pushes them across the threshold. In other words, the price decrease prior to the thresholds is explained because the incentive to lower prices from the *induced review effect* is amplified due to a discontinuous increase in the future value of the rating stock when the threshold can be crossed.

A.2 Quality-Learning Variant

In the quality-learning model, we implement the same setting as described above with the modification that hosts also update their beliefs about the apartment's quality. Their updating, however, is dynamic in that they update after each individual review using the same rule as consumers and enter the next period with a new belief. We solve the model accordingly in the same way but with an additional state variable,

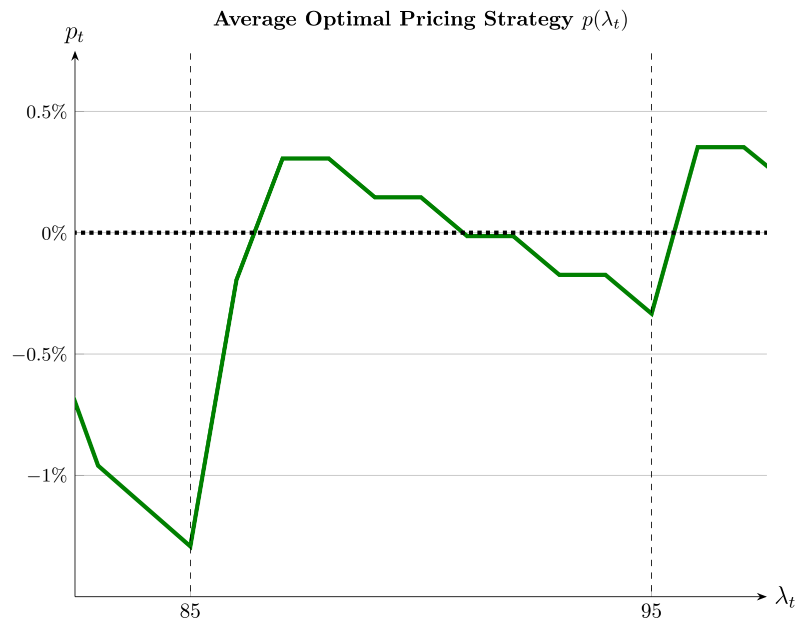

Figure 7: Optimal Pricing around Thresholds by Strategic Hosts — Model Prediction

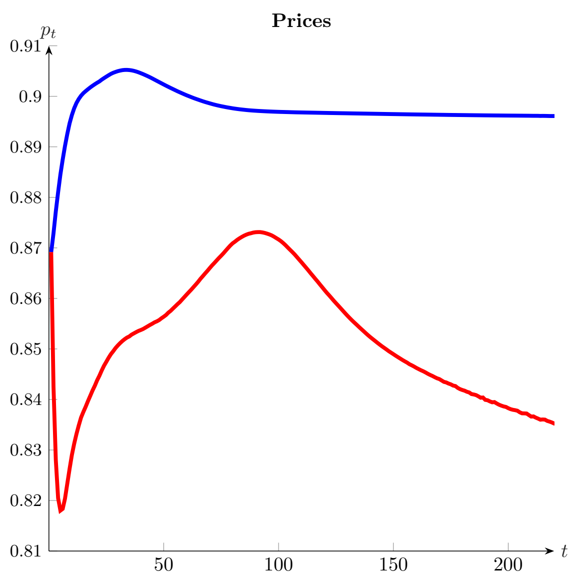

Figure 8: Quality-learning hypothesis: price paths high (blue) vs. low (red) quality

which is the host's belief.

The plots are generated using the optimal strategy and simulating the review 250,000 times for each realized but unobserved quality. When we average the results, we average them according to the prior belief μ_0. The results are displayed in Figure 5. Figure 8 shows the absolute price paths and verifies that both low and high-quality listings exhibit an upward price trend after the initial periods. However, the upward trend is much more immediate for high-quality listings as low-quality listings exhibit a downward adjustment upon inferring their low quality from initial reviews.

Overall, this implies that it is high-quality listings that would be classified as having a first-period discount according to our discount measure, while low-quality listings would be measured as exhibiting little to no discount on average.

A.3 Optimal Pricing around Thresholds

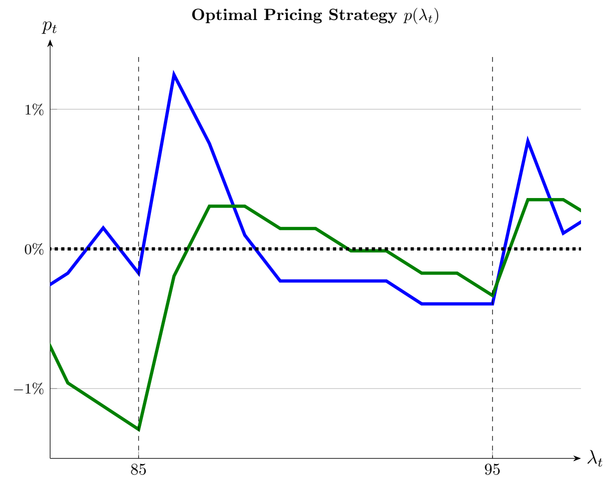

Figure 9: Optimal Pricing around Thresholds by Strategic Hosts — Comparison

While the main specification above assumes a dominant value-for-money effect in line with the empirical findings, we can also numerically implement a dominant selection effect. We contrast the implications for optimal pricing by strategic hosts around rating thresholds in Figure 9. It plots the average optimal pricing strategy around the latent thresholds of 85 and 95, and contrasts the optimal strategy with a dominant value-for-money effect (blue) and that with a dominant selection effect (red). Because the rating update and dominant selection effect work in opposite directions below the threshold, the price decrease is muted, while the increase after passing the threshold is exacerbated.

B Aggregation of Individual Ratings

Figure 10 shows the probability of observing a change in the aggregate overall rating between two consecutive monthly updates as a function of the number of ratings a listing has already received. The continuously decreasing probability of observing aggregate rating changes is consistent with simple averaging of individual ratings, which is also the specification of the updating process in our model. It directly follows from the diminishing probability of rating changes that any effect of prices on ratings should be less pronounced for listings that had received more ratings in the past.

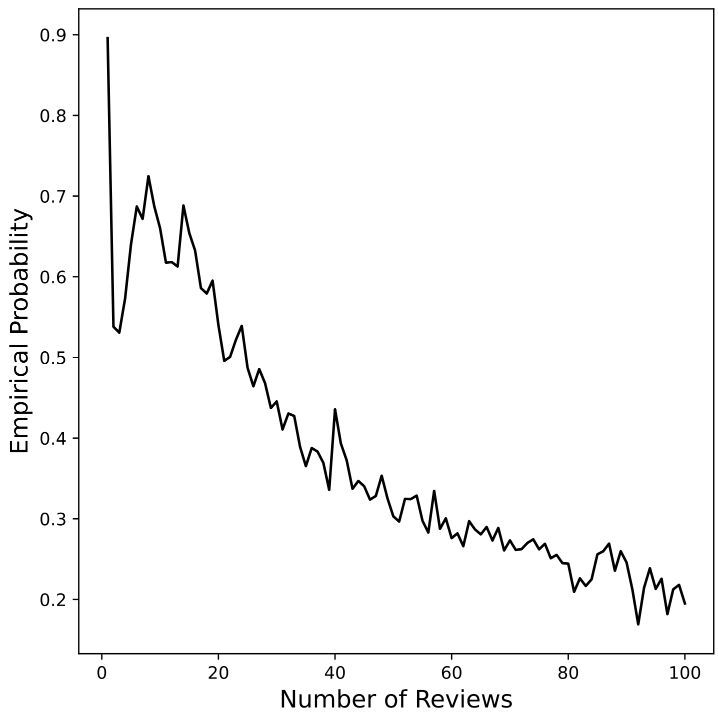

Figure 10: Empirical Probability of Observing a Rating Change

Notes: The empirical probability is obtained by first dropping all observations for which we observe no change in the number of ratings between two consecutive scrapes in the InsideAirbnb data. We then divide the number of observed ratings changes by the total number of observations for each observed number of ratings. We only display the results for observations with 100 or fewer ratings.

C Impact of Ratings on Host Revenues

Our model assumes that ratings positively affect profits. Otherwise, sellers would have no incentive to strategically influence ratings. Past research found evidence in line with this assumption (see Section 4 of Tadelis, 2016, for an extensive review). In this appendix, we analyze whether our data also support this assumption.

For our analysis, we exploit the granular measure available for the overall rating. This measure ranges on a scale from 20 to 100 in increments of one unit and was not observed by travelers at the time our data were sampled. Instead, travelers only observed a less granular rating ranging from 1 to 5 stars in increments of half a star.

The number of stars shown to customers is a step function of the underlying granular rating measure: If we denote by r the granular measure, the number of stars, $f(r)$, shown to a potential traveler visiting the listing webpage on Airbnb depends on the interval r lies in. For example $f(r) = 1$ if $r \in [20, 25)$, $f(r) = 1.5$ if $r \in [25, 35)$, $f(r) = 2$ if $r \in [35, 45)$, $\cdots$, $f(r) = 4.5$ if $r \in [85, 95)$, and $f(r) = 5$ if $r \in [95, 100]$.

We exploit the discontinuities in the relationship between the salient star rating and the granular measure to implement a regression discontinuity design by running the following regression:

$$y_{it} = \beta_0 + \beta_1 \mathcal{I}_{bw}(r_{it} > \tau) + X'_{it}\gamma + \mu_i + \mu_t + \epsilon_{it}, \tag{14}$$

where y_{it} denotes the revenue in euros, and X_{it} contains the number of days a listing is available, the logarithm of the offer price, the logarithm of the host response rate, and an indicator variable for the deciles of the number of ratings. μ_i denotes the listing fixed effects, and μ_t denotes the month fixed effects. $\mathcal{I}_{bw}(r_{it} > \tau)$ is an indicator variable that takes the value one if the non-salient rating r at the *beginning* of period t exceeds a specified threshold, τ, within the closed interval with bandwidth bw. For example, if $bw = 0.5$ and $\tau = 80.5$, then β_1 captures the conditional average difference in revenues between listings with a rating of 81 or 80.

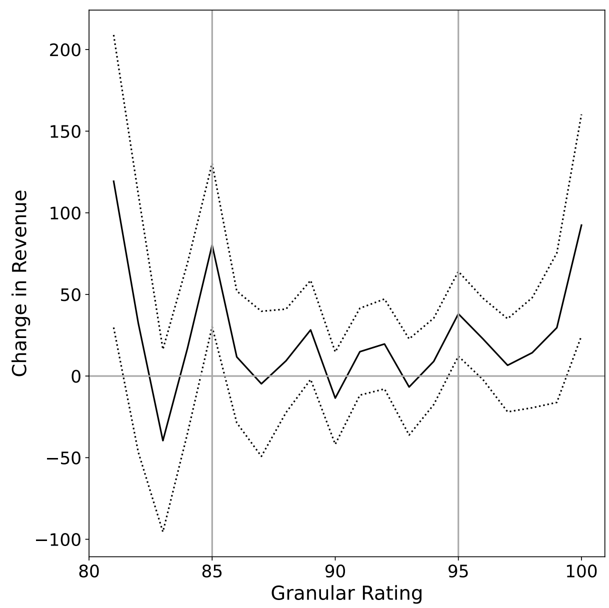

Figure 11: Impact of Incremental Granular Rating Change Using All Listings

Notes: The 95% confidence intervals are based on heteroskedasticity-robust standard errors.

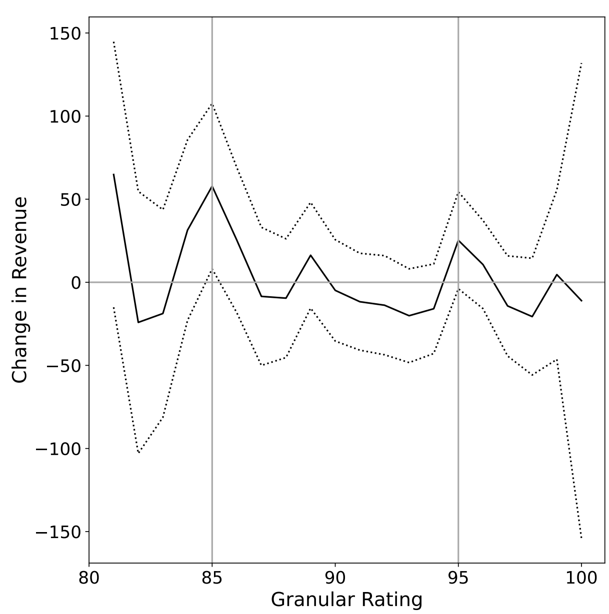

Figure 12: Impact of Incremental Granular Rating Change Using Only Listings that Experience Rating Changes

Notes: The 95% confidence intervals are based on heteroskedasticity-robust standard errors.

In the following, we start by fixing $bw = 0.5$ and let $\tau = \{80.5, 81.5, 82.5, ..., 99.5\}$. Only two thresholds result in a salient rating change. We denote salient thresholds by $\tau^\dagger$. The first salient threshold ($\tau^\dagger = 84.5$) marks the transition from four to four-and-a-half stars and the second salient threshold ($\tau^\dagger = 94.5$) marks the transition from four-and-a-half to five stars.

Figure 11 shows the estimators for β_1 for the different values of the threshold and the respective 95% confidence intervals. The vertical lines indicate the salient thresholds. The results for the non-salient thresholds constitute placebo tests. The coefficients indicate statistically significant effects at both salient thresholds. However, it is also apparent that we detect statistically significant effects at some non-salient thresholds.

In Figure 12, we refine the analysis of Figure 11 by estimating Equation (14) only for listings for which we observe changes in the granular rating measure. By doing so, we enforce that the listings on both sides of a threshold are the same. This allows us to address concerns related to unobserved type heterogeneity of listings on both sides of a threshold. To capture the immediate effect of a rating change, we focus on the observations in the periods immediately before and after a granular rating change occurs. Figure 12 shows a statistically significant effect for the salient threshold from four to four-and-half stars. The measured effect from four-and-half to five stars is not significant at conventional levels.

To address potential power issues, we relax the analysis shown in Figure 12 in two dimensions: First, we use all observations before and after a rating change (not only the observations *immediately* before and after); second, we vary the size of the bandwidth. The results for the thresholds $\tau = \{84.5, ..., 94.5\}$ and the various bandwidths, bw, are shown in Table 6.

Note that for bandwidths larger than 0.5, we exclude some thresholds to run certain placebo tests: for example, with a non-salient threshold of 85.5 and a bandwidth of 1.5, a listing changing its granular rating from 84 to 86 crosses both the salient threshold of 84.5 and the non-salient threshold of 85.5. To avoid having "double-crossing" affect our analysis, we only use a non-salient threshold, τ, for a placebo-test if $\tau^\dagger \leq \tau - bw$ and $\tau^\dagger \geq \tau + bw$.

Table 6 shows statistically significant effects for both salient thresholds. The estimated magnitudes are relatively stable and do not appear to change systematically with the size of the bandwidths. For the first salient threshold (four to four-and-a-half stars), we obtain a point estimate of approximately 70 euros. For the second salient threshold (four-and-a-half to five stars), we obtain a point estimate of approximately 42 euros.

Table 6: Results of Regression Discontinuity Analysis

	Thresholds										
bw	$84.5^{\dagger}$	85.5	86.5	87.5	88.5	89.5	90.5	91.5	92.5	93.5	$94.5^{\dagger}$
0.5	94.21^{***}	33.88	9.08	−18.68	19.28	-39.17^{*}	7.08	29.01	7.44	25.68	49.68^{**}
	(27.57)	(26.24)	(34.74)	(20.63)	(20.21)	(19.70)	(19.10)	(18.71)	(21.92)	(18.55)	(18.95)
1.5	66.03^{**}		20.05	10.93	15.86	−2.44	4.55	3.15	12.48		38.89^{**}
	(22.13)		(21.62)	(17.30)	(16.23)	(17.41)	(14.64)	(13.40)	(18.58)		(14.02)
2.5	71.25^{***}			17.93	26.29	9.95	8.51	11.56			42.71^{***}
	(19.52)			(14.53)	(15.19)	(15.46)	(13.53)	(12.86)			(12.93)
3.5	75.52^{***}				29.97^{*}	21.25	16.04				41.87^{***}
	(21.79)				(14.50)	(15.55)	(13.11)				(12.59)

Notes: *, **, *** indicate statistical significance at the five, one, and 0.1 percent levels, respectively. The heteroskedasticity-robust standard errors are in parentheses. Daggers denote the salient thresholds. Thresholds from 85.5 to 93.5 are used for the placebo tests. For each bandwidth, we only use non-salient thresholds, τ, for the placebo tests if $\tau^{\dagger} \leq \tau - bw$ and $\tau^{\dagger} \geq \tau + bw$.

D Robustness of Price-Rating Regressions

Table 7: First-Differences Regression of Ratings

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
All Listings							
log(price)	-0.028^{***}	-0.039^{***}	-0.001	-0.021^{***}	-0.008	-0.027^{***}	-0.009^{*}
	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
Split by Price Level							
log(price) - Low	-0.059^{***}	-0.065^{***}	-0.005	-0.039^{***}	-0.024^{*}	-0.034^{***}	-0.002
	(0.007)	(0.011)	(0.008)	(0.010)	(0.011)	(0.009)	(0.009)
log(price) - High	-0.012^{*}	-0.026^{***}.	0.001	-0.012	-0.001	-0.023^{***}	-0.012^{*}
	(0.005)	(0.007)	(0.006)	(0.007)	(0.008)	(0.006)	(0.006)
Obs.	134,946	134,815	134,825	134,905	134,942	134,879	134,838

Notes: $*$, $**$, $***$ indicate statistical significance at the five, one, and 0.1 percent levels, respectively. Standard errors are shown in parentheses and are clustered at the listing level. Each regression additionally controls for the month fixed effects, a second-order polynomial of the number of ratings, and the host response rate as a proxy for the effort of the host. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication rating, respectively.

E Pricing Patterns Around Salient Rating Thresholds

In this appendix, we provide additional information on the estimation of the quality fixed effects used in Equation (11). Additionally, we provide robustness checks for the results presented in Figure 3 when including further observable characteristics of the listings.

To create the quality fixed effects, we first run the following regression:

$$p_{it} = \mu_i + \mu_t + \epsilon_{it}, \tag{15}$$

where p_{it} denotes the offer price of listing i in month t, μ_i captures the listing fixed effects, and μ_t the month fixed effects. After estimating Equation (15), we obtain the quintiles of μ_i based on which we create the five-scale indicator variable capturing the quality of a given listing i.

Figure 13 repeats the analysis shown in Figure 3 but additionally controls for the room type (apartment, private room, or shared room), the number of bedrooms and bathrooms, and the logarithm of the host response rate. Figure 14 presents the results obtained when pooling observations across both thresholds and measuring the conditional average price as a function of the distance to the closest salient threshold.

Using Listing Fixed-Effects Instead of Quality Fixed-Effects

Instead of quality fixed effects, a natural approach would be to use listing fixed effects. An analysis based on listing fixed effects is complicated because the within-listing variation is relatively scant. For example, in the present analysis, more than 60 percent of listings close to the threshold from four to four-and-a-half stars experience no more than one rating change.[37] The situation is similar at the upper threshold, with more than 60 percent experiencing no more than one change.

This lack of listings with a "longer" trajectory across different ratings close to the thresholds frustrates a non-parametric measurement of the pricing behavior around the thresholds. Running a listing fixed-effects regression analogous to Equation (11) over a certain range of the granular rating measure does not reveal a clear pattern around the salient thresholds. By contrast, using the quality indicator variable allows us to exploit the cross-sectional variation of listings similarly valued by consumers (quality fixed effects) and in close geographical proximity (location fixed effects).

[37] We here define a listing as close to the threshold if its granular overall rating is within three units of the threshold.

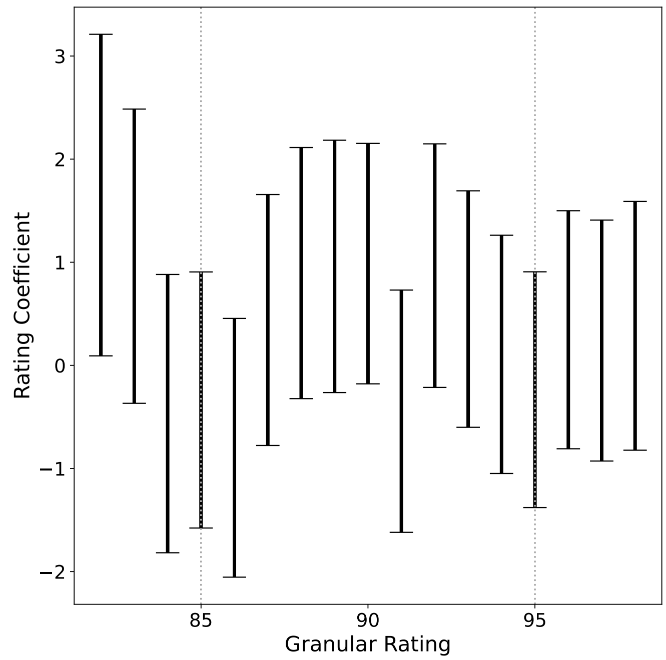

Figure 13: Conditional Average Offer Prices Around Salient Threshold

Notes: The 95% confidence intervals are based on heteroskedasticity-robust standard errors.

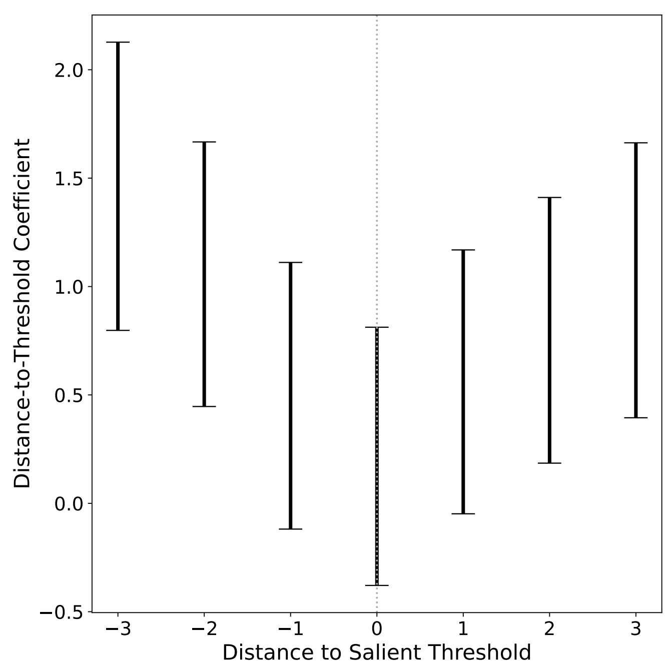

Figure 14: Conditional Average Offer Prices Around Salient Thresholds (When Pooling Thresholds)

Notes: The 95% confidence intervals are based on heteroskedasticity-robust standard errors.

Regression Tables

Table 8: Average Offer Price Around Salient Threshold

	Difference to Omitted Category (Overall rating = 81)
Overall Rating = 82	1.388
	(0.767)
Overall Rating = 83	0.337
	(0.765)
Overall Rating = 84	−0.954
	(0.693)
Overall Rating = 85	−1.432*
	(0.630)
Overall Rating = 86	−1.254*
	(0.635)
Overall Rating = 87	−0.261
	(0.620)
Overall Rating = 88	0.197
	(0.620)
Overall Rating = 89	0.627
	(0.624)
Overall Rating = 90	0.964
	(0.586)
Overall Rating = 91	−0.268
	(0.592)
Overall Rating = 92	1.034
	(0.598)
Overall Rating = 93	0.666
	(0.578)
Overall Rating = 94	0.262
	(0.581)
Overall Rating = 95	0.000
	(0.572)
Overall Rating = 96	0.576
	(0.581)
Overall Rating = 97	1.052
	(0.589)
Overall Rating = 98	1.112
	(0.614)
log(host response rate)	0.246
	(0.254)
Num.Obs.	156,634
R2 Adj.	0.698

Notes: * indicates statistical significance at the five percent level. The regression additionally controls for location fixed effects. The heteroskedasticity-robust standard errors are in parentheses.

F Price-Rating Dynamics

Regression Tables

Table 9: Results for Rating Categories

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
Disc.	0.334^{***}	0.415^{***}	0.100	0.114	0.370^{***}	0.104	0.140
	(0.101)	(0.106)	(0.072)	(0.093)	(0.105)	(0.097)	(0.097)
$\beta_1\times \mathcal{I}(\lambda(i,t)=2)$	-0.072	-0.022	0.049	-0.022	-0.055	-0.038	0.009
	(0.050)	(0.055)	(0.046)	(0.054)	(0.060)	(0.046)	(0.049)
$\beta_2\times \mathcal{I}(\lambda(i,t)=3)$	-0.147^{*}	-0.111	-0.047	-0.048	-0.075	-0.048	-0.010
	(0.065)	(0.069)	(0.056)	(0.071)	(0.078)	(0.060)	(0.059)
$\beta_3\times \mathcal{I}(\lambda(i,t)=4)$	-0.200^{**}	-0.218^{**}	-0.064	-0.040	-0.147	-0.045	-0.018
	(0.072)	(0.075)	(0.061)	(0.075)	(0.083)	(0.065)	(0.066)
$\beta_4\times \mathcal{I}(\lambda(i,t)=5)$	-0.243^{**}	-0.259^{**}	-0.079	-0.074	-0.104	-0.045	-0.074
	(0.080)	(0.081)	(0.063)	(0.080)	(0.086)	(0.076)	(0.074)
$\beta_5\times \mathcal{I}(\lambda(i,t)=6)$	-0.259^{**}	-0.297^{***}	-0.071	-0.129	-0.100	-0.066	-0.096
	(0.080)	(0.084)	(0.064)	(0.079)	(0.087)	(0.079)	(0.076)
R2 Adj.	0.089	0.076	0.239	0.080	0.091	0.077	0.082
Num.Obs.	44,616	44,435	44,430	44,572	44,615	44,550	44,461

Notes: $*$, $**$, $***$ indicate statistical significance at the five, one, and 0.1 percent levels, respectively. Standard errors are shown in parentheses and are clustered at the listing level. All regressions control for listing location fixed effects, month fixed effects, and month-of-market-entry fixed effects. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively. In contrast to Figure 4, the interactions between the period dummies and the discount capture the change of the the impact of the discount on the variable of interest from period $\lambda(i,t)=2$ to $\lambda(i,t)=6$ relative to the baseline period (Disc.).

Table 10: Results for Performance Measures

	Price	log(Price)	Bookings	log(Bookings)	Revenue	log(Revenue)
Disc.	-67.057^{***}	-0.713^{***}	2.206^{***}	0.714^{***}	-251.697^{*}	-0.068
	(5.415)	(0.049)	(0.196)	(0.059)	(114.612)	(0.077)
$\beta_2 \times \mathcal{I}(\lambda(i,t)=2)$	56.484^{***}	0.654^{***}	-0.378	-0.165^{**}	669.621^{***}	0.584^{***}
	(1.970)	(0.014)	(0.195)	(0.062)	(87.964)	(0.072)
$\beta_3 \times \mathcal{I}(\lambda(i,t)=3)$	83.66^{***}	0.964^{***}	-0.788^{***}	-0.331^{***}	1213.085^{***}	0.660^{***}
	(2.718)	(0.017)	(0.221)	(0.066)	(112.334)	(0.077)
$\beta_4 \times \mathcal{I}(\lambda(i,t)=4)$	97.61^{***}	1.126^{***}	-1.492^{***}	-0.520^{***}	1304.773^{***}	0.614^{***}
	(2.667)	(0.020)	(0.233)	(0.070)	(116.073)	(0.077)
$\beta_5 \times \mathcal{I}(\lambda(i,t)=5)$	110.182^{***}	1.228^{***}	-2.148^{***}	-0.676^{***}	1263.001^{***}	0.552^{***}
	(3.375)	(0.023)	(0.252)	(0.075)	(120.546)	(0.078)
$\beta_6 \times \mathcal{I}(\lambda(i,t)=6)$	110.753^{***}	1.233^{***}	-2.526^{***}	-0.772^{***}	1221.145^{***}	0.444^{***}
	(3.374)	(0.025)	(0.262)	(0.074)	(122.529)	(0.082)
$\log(\text{reviews})_{t-1}$	-1.708^{***}	-0.016^{***}	0.159^{***}	0.035^{***}	-6.492^{*}	0.003^{*}
	(0.123)	(0.001)	(0.005)	(0.001)	(2.548)	(0.001)
Days Avail.	0.334^{***}	0.004^{***}	0.114^{***}	0.036^{***}	60.279^{***}	0.055^{***}
	(0.041)	(0.000)	(0.001)	(0.000)	(0.757)	(0.001)
R2 Adj.	0.310	0.362	0.325	0.347	0.349	0.454
Num.Obs.	48,756	48,750	48,756	48,756	48,756	48,756

Notes: $*$, $**$, $***$ indicate statistical significance at the five, one, and 0.1 percent levels, respectively. Standard errors are shown in parentheses and are clustered at the listing level. All regressions control for listing location fixed effects, month fixed effects, and month-of-market-entry fixed effects. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively. In contrast to Figure 4, the interactions between the period dummies and the discount capture the change of the the impact of the discount on the variable of interest from period $\lambda(i,t) = 2$ to $\lambda(i,t) = 6$ relative to the baseline period (Disc.).

Including Listing Fixed Effects

Because d_i does not vary within listings, a fixed effect estimation of Equation (12) can only identify *changes* in the relationship between y_{it} and d_i across periods – but not the relationship between y_{it} and d_i in the baseline period (which we define as the first period). Figure 15 shows the result of an analysis with listing fixed effects for the dependent variables presented in the main text. As can be seen, the estimated β_τ are consistent with the results of the main analysis. In other words, the changes shown in Figure 15 correspond well to the changes observed in Figure 4 between the coefficients for $\tau \geq 2$ and $\tau = 1$.

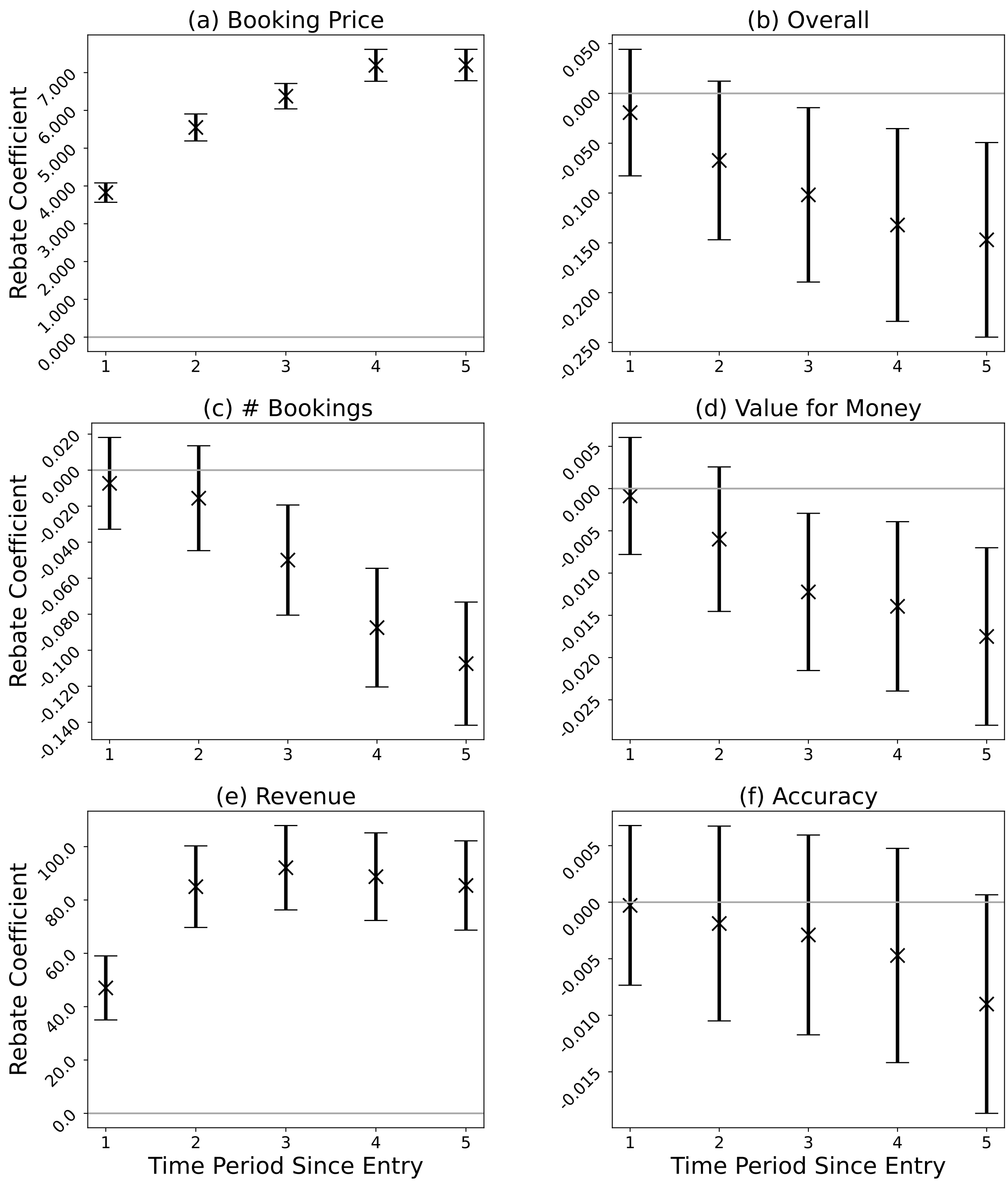

Figure 15: Price-Rating Dynamics – Including Listing Fixed-Effects

Notes: The results are obtained from regressions including listing fixed effects. The 95% confidence intervals are based on standard errors clustered at the listing level.

G Using Demand as an Instrumental Variable

In this subsection, we present the results we obtain when estimating Equation (12) using two demand-based variables as excluded instruments for the discount factor. The first instrument is the ratio between the daily average number of bookings within a radius of 250 meters in the month of entry and the average of this same variable in the subsequent months. This can be considered as the demand-equivalent of the price discount variable defined in Equation (13). The demand within a radius of 250 meters in a given month is calculated based only on the days in which the focal listing was booked. The second instrument is the ratio between the daily average number of bookings observed within 250 meters of each listing and the average value of this variable for all other listings in their month of entry. The variable captures the local demand conditions experienced by a listing relative to the demand conditions experienced by other listings when they entered the market. Both instruments are interacted with a set of dummy-variables that count the number of months since a listing entered the market.

Our instruments are valid if the timing of entry is uncorrelated with unobserved factors that influence both the discount and ratings. One might, for example, be concerned that more sophisticated hosts time their entry decision strategically, or that variations in demand coincide with changes in consumer types and how these consumers rate specific characteristics. While we cannot rule out these possibilities entirely, we emphasize that over-identification tests predominantly do not reject the null-hypothesis of instrument validity when estimating the effect of discounts on ratings (see Appendix G). Finally, we also note that the observed relationship between discounts and ratings is inconsistent with a model in which discounts are driven by unobserved listing quality of the type discussed in Section 5.4.

The estimated coefficients in Figure 16 are an order of magnitude larger than the coefficients obtained by standard *OLS* and less precisely estimated (the detailed regression results can be found in Table 11 in Appendix G). Overall, the results confirm the insights of the *OLS* analysis insofar as the trends observed for the overall and the value-for-money rating are both downward sloping. Interestingly, the location rating now displays an upward trend, which is consistent with a dominant selection effect for this rating category (Table 11 shows that the difference between the first period coefficients and the coefficients in other periods are, in fact, mostly significant). Note that the demand instrument can not be used to analyze the evolution of bookings over time because the exclusion restriction is likely to be violated.[38]

[38] The over-identification test rejects the null-hypothesis of joint instrument validity for the variables related to demand. By contrast, among all rating variables, the null-hypothesis of instrument validity is only marginally rejected for the location rating. The first-stage F statistics and the results for Sargan's J test for the over-identifying restrictions are presented in Appendix G.

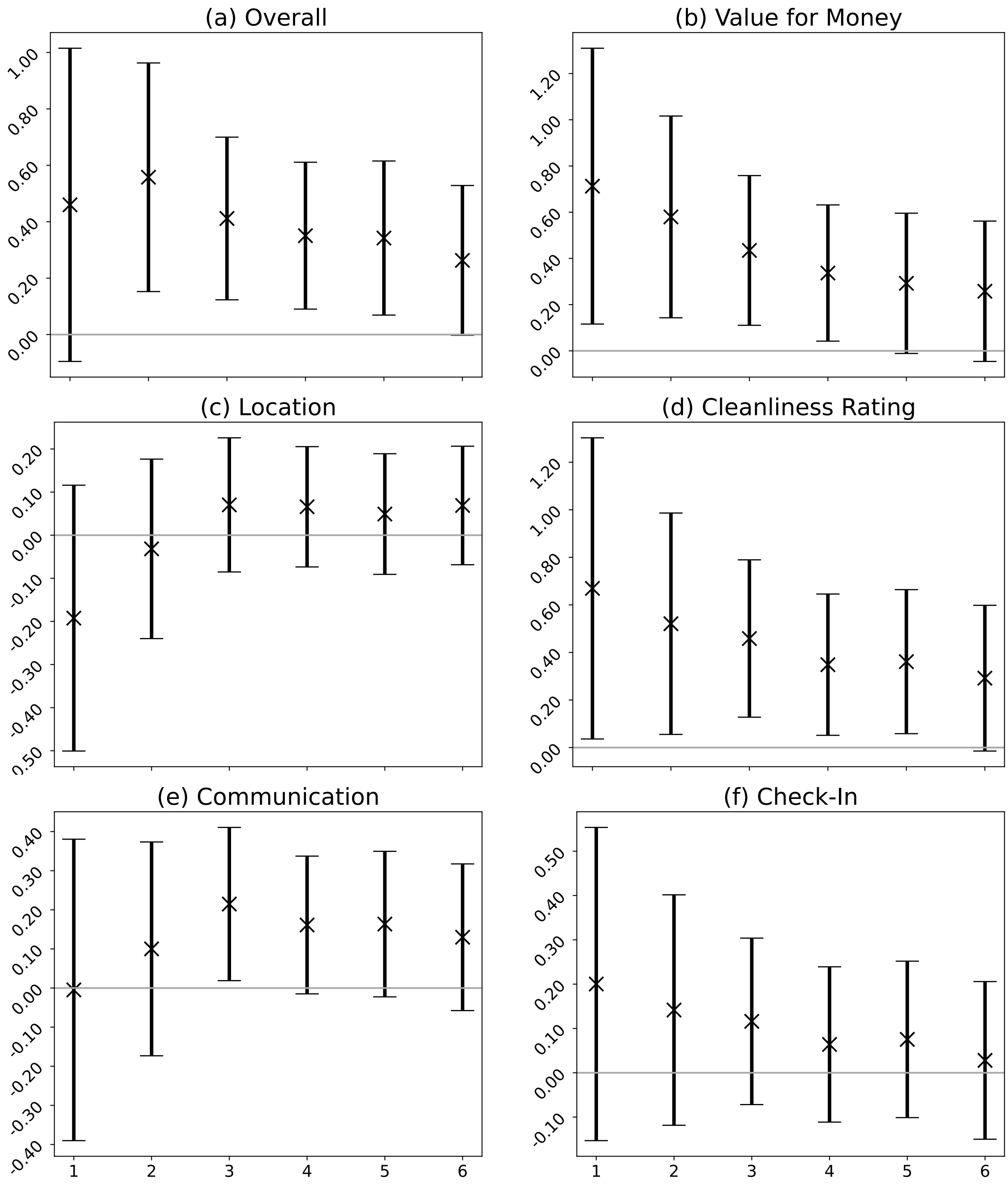

Figure 16: $2SLS$-Results Scaled at Median Discount

Notes: The 95% confidence intervals are based on standard errors clustered at the listing level.

Table 11: IV Results for Rating Categories

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
$\beta_1 \times \mathcal{I}(\lambda(i,t)=1)$	6.536	10.162*	−2.653	6.940*	9.124*	0.075	3.009
	(4.025)	(4.336)	(2.241)	(3.518)	(4.562)	(2.804)	(2.575)
$\beta_2 \times \mathcal{I}(\lambda(i,t)=2)$	7.925**	8.339**	−0.264	5.638*	7.178*	1.686	2.186
	(2.938)	(3.160)	(1.509)	(2.497)	(3.339)	(1.991)	(1.891)
$\beta_3 \times \mathcal{I}(\lambda(i,t)=3)$	5.848**	6.311**	1.171	3.734*	6.435**	3.273*	1.811
	(2.090)	(2.328)	(1.119)	(1.826)	(2.347)	(1.421)	(1.357)
$\beta_4 \times \mathcal{I}(\lambda(i,t)=4)$	4.981**	4.961*	1.125	3.394*	4.994*	2.517	1.149
	(1.888)	(2.141)	(1.011)	(1.637)	(2.125)	(1.292)	(1.279)
$\beta_5 \times \mathcal{I}(\lambda(i,t)=5)$	4.862*	4.280	0.739	2.967	5.472	2.455*	1.176
	(1.980)	(2.258)	(1.029)	(1.736)	(2.192)	(1.402)	(1.320)
$\beta_6 \times \mathcal{I}(\lambda(i,t)=6)$	3.736	3.615	1.045	2.087	4.089	1.818	0.344
	(1.925)	(2.174)	(0.980)	(1.697)	(2.175)	(1.342)	(1.276)
Num.Obs.	44,616	44,435	44,430	44,572	44,615	44,550	44,461

Notes: $*$ indicates statistical significance at the five percent level, respectively. Standard errors are shown in parentheses and are clustered at the listing level. All regressions control for listing location fixed effects, month fixed effects, and month-of-market-entry fixed effects. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively.

Table 12: First Stage F statistics – $2SLS$ Entry Price Regression

	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 6$
F stat	6.86	10.51	12.52	12.51	11.64	11.95
p-val	0.00	0.00	0.00	0.00	0.00	0.00

The table reports the first stage F-statistics for the various endogenous regressors of the $2SLS$ estimator used in Figure 16. Note that the first stage is identical for each dependent variable. The endogenous regressors correspond to the interaction terms between the initial discount and the period fixed-effect τ.

Table 13: Sargan's J Test for Various Dependent Variables

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-In	Days B.	# B.
Stat	4.48	10.28	13.12	2.32	0.79	11.18	10.28	39.37	72.89
p-val	0.61	0.11	0.04	0.89	0.99	0.08	0.11	0.00	0.00

The table reports the test statistic and the p-value of Sargan's J test. The Null-Hypothesis that the over-identifying restrictions are valid cannot be rejected for the equations with the rating statistics as dependent variables. The Null-Hypothesis is rejected when the dependent variable is the booking price or the number of bookings. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication rating, respectively. Days B. denotes the days booked, and # B. the number of bookings.