Value for Money and Selection: How Pricing Affects Airbnb Ratings

Christoph Carnehl
Maximilian Schaefer
André Stenzel
Kevin Ducbao Tran

Discussion Paper 22/771

31 August 2022

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of Bristol
School of Economics

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

Christoph Carnehl†

Maximilian Schaefer‡

André Stenzel§

Kevin Ducbao Tran¶

August 31, 2022

Abstract

We investigate the impact of prices on ratings using Airbnb data. We theoretically illustrate two opposing channels: higher prices reduce the value for money, worsening ratings, but they increase the taste-based valuation of the average traveler, improving ratings. Results from panel regressions and a regression discontinuity design suggest a dominant value-for-money effect. In line with our model, hosts strategically complement lower prices with higher effort more when ratings are relatively low. Finally, we provide evidence that, upon entry, strategic hosts exploit the dominant value-for-money effect. The median entry discount of seven percent improves medium-run monthly revenues by three percent.

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, Northwestern University, and Yale University, 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, as well as 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

‡Yale University and University of Bologna, maximilian.schaefer@yale.edu

§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 Germany Science 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 platforms such as Airbnb, eBay, and Deliveroo match buyers and sellers who typically have not previously transacted with each other. As the goods or services cannot be physically examined pre-purchase, these platforms seek to build and foster trust between market participants to resolve inherent quality uncertainty. A key instrument for this purpose is the use of reputation and feedback systems (Tadelis, 2016), which are prevalent on virtually every online platform. Ideally, these systems 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 for designing effective reputation and feedback systems.

In this paper, we derive and empirically assess predictions regarding hosts' strategic pricing incentives when prices affect ratings. 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 ratings. First, we show that, all else equal, higher prices are associated with lower ratings. Second, consistent with our theoretical analysis, we provide evidence that hosts charging relatively lower prices when entering the platform benefit from such early discounts in the medium run. A discount of five euros per night (seven percent) in the first period 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 price of the listing and an effort level that affects the travelers' experience; for example, travelers may benefit from a smoother check-in experience or better communication. Prospective travelers observe two types of aggregate ratings: a value-for-money rating and an effort rating. Based on these ratings, travelers form expectations about the quality of the listing and the host's effort. Travelers' ultimate booking decisions depend on these expectations, their idiosyncratic preferences for the listing, and its price.

Travelers who stay at an apartment provide ratings that depend on the apartment's quality, the host's effort, the travelers' idiosyncratic tastes, and the price. We assume that higher quality, effort, and taste induce higher value-for-money ratings. Further, we assume that a higher price lowers the value-for-money ratings directly. However, a higher price could also increase the ratings indirectly because the average traveler who decides to book the apartment will have a higher idiosyncratic taste: following a price increase, the previously marginal traveler, who had the lowest taste of all purchasing consumers, will no longer rent the apartment. The first effect is what we refer to as the *value-for-money effect*, while the latter effect is what we refer to as the *selection effect*.

The net effect of prices on ratings depends on the relative importance of the value-for-money and selection effects. For the effort rating, we assume that higher effort levels lead to better ratings while higher prices have weakly negative impacts. Importantly, the host's current pricing and effort decisions will influence the ratings posted by current-period travelers and, thereby, 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 scores, as well as the optimal price and effort levels. We discuss the insights of our model along with our empirical findings below.

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 ratings on lagged prices and controlling for listing-specific fixed effects, we find that higher prices are generally associated with lower ratings. This relationship is particularly pronounced for value-for-money ratings and disappears for location ratings. While the negative relationship between prices and value-for-money 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 weaker value-for-money and a 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.

When analyzing hosts' pricing behaviors in relation to salient ratings thresholds, we find evidence that hosts are aware of the dominant value-for-money effect. Hosts who are close to but below the threshold above which their displayed star rating improves tend to lower their prices compared to hosts further away from the threshold. Past these rating thresholds, prices tend to be higher, suggesting that hosts profit from the improved ratings.

Decomposing the analysis with respect to above- and below-median-priced listings, we find a stronger negative relationship between prices and ratings for lower-priced listings. This result suggests a higher price sensitivity in ratings among travelers who choose cheaper accommodations. In the context of our model, this higher price sensitivity increases the relative importance of the value-for-money effect.

Within our model, a host's effort and pricing decisions can be either strategic complements or substitutes in managing their ratings. The shape of the host's continuation profits as a function of the rating determines the strategic interaction between these two factors. The shape, in turn, may vary with the level of the rating. In particular, if a rating increase leads to a higher marginal benefit of further rating increases—that is, if future profits are convex in the rating—price and effort are strategic complements. Our empirical analysis establishes this convex relationship for lower ratings while it becomes concave as ratings increase.[1] Our model, therefore, predicts that effort and prices are complements when a listing has lower ratings and that effort and prices are substitutes when a listing has higher ratings. Regressing

[1] The documented shape is intuitive from the perspective of (Bayesian) belief adjustments. An upward adjustment due to a positive signal would likely be stronger for an intermediate belief than when the traveler is almost certain that the apartment is of higher value.

effort ratings on prices, we indeed find a negative relationship for listings with low overall ratings but that this relationship becomes less negative and even positive for higher ratings.

Finally, our theoretical analysis suggests that hosts can use strategic entry pricing, which sacrifices short-run profits, to obtain better ratings and increase profits in subsequent periods. The underlying reason is that prices can be adjusted frequently but also have a persistent effect on the ratings, which determine future demand. Given that the value-for-money effect overall dominates, hosts should offer a price discount when entering the market relative to a naive entry price that ignores this dynamic effect.

This theoretical hypothesis is again supported empirically. We find that new hosts who set a discount when entering the market receive better value-for-money ratings and more bookings in the early periods. This allows these hosts to set relatively higher prices than other listings in subsequent periods without offering a quantity reduction, resulting in higher revenues in the medium run. A discount of five euros (seven percent) in the first-period price increases medium-run monthly revenues by approximately 68 euros (three percent) compared to listings that do not offer a discount when entering the market.

Related literature. The literature on the ratings-prices nexus has mostly focused on how ratings affect prices. Several studies establish a robust positive relationship between ratings and prices (Teubner et al., 2017), revenues (Luca, 2016), and quantities (Livingston, 2005).

In contrast, we are interested in the opposite mechanism: whether prices affect ratings. We add 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. (2022), albeit with a different focus. They study the theoretical long-run properties of learning, ratings, and prices, as well as the rating system's design. In contrast, our model illustrates the effects at play and adds an effort component to the host's decision. This theoretical framework also helps reconcile the results in the empirical literature. Closely related is concurrent work by Jeziorski and Michelidaki (2021), who relate different sub-categories of ratings to different frames and empirically show that the explicit value-for-money frame features the highest degree of price responsiveness. 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 the paper by considering continuous variation in prices rather than comparing a price of zero and positive prices.[2] Furthermore, we consider an additional effect that prices can have on ratings: the value-for-money effect, which seems to dominate in our data. 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. Our theoretical model explains this result in the form of a dominant value-for-money effect. Sorokin (2021) finds that producers on the video game platform Steam use discounts to transition to higher ratings tiers. Our paper contributes to this literature by proposing a unified theoretical framework to explain these empirical results. Our empirical part moreover provides evidence in line with this framework.

Our paper is also more broadly related to research on the determinants of ratings. Cabral and Li (2015) find that lower quality transactions result in more negative feedback. Mayzlin et al. (2014) find evidence of hotels faking negative ratings for their competitors and positive ratings for themselves on TripAdvisor. Luca and Zervas (2016) provide evidence of restaurants using fake ratings on Yelp and He et al. (2022) study the market for fake product ratings on Amazon.com and show that these seem to be used mostly for low quality products. Proserpio and Zervas (2017) find that when responding to ratings, hotels tend to receive fewer but longer negative ratings. We add to this literature by highlighting theoretically and empirically how price setting can affect ratings.

We proceed as follows. In Section 2, we describe our theoretical framework and derive some testable hypotheses for the empirical part. In Section 3, we introduce our data and provide some descriptive statistics. We describe our main empirical analysis and results in Section 4. In Section 5, we focus on analyzing strategic price setting by hosts entering the market. Finally, Section 6 concludes.

2 Model

We begin by setting up our theoretical framework before deriving testable hypotheses for our empirical analysis. Our model consists of a single host offering an apartment to a continuum of travelers. 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.[3] We outline each component of the model in more detail below.

[2]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).

[3]For further details, see Carnehl et al. (2022), who explicitly analyze extensions to competitive settings in a related model.

Host There is a single long-lived host with an apartment of fixed quality, $\theta \in \{L, H\}$. The apartment has an initial rating of $\Psi_0 \in \mathbb{R}^2_+$, which parameterizes the initial attitude of potential travelers towards the apartment; that is, Ψ_0 serves the role of a prior about the apartment's quality. In each period $t = 1, \ldots, T$ with $T \leq \infty$, the host chooses a price $p \geq 0$ and effort $e \geq 0$ at cost $c(e) = \frac{c}{2}e^2$. The host maximizes their discounted total profits and has a discount factor of $\delta \in [0, 1)$.

Observed ratings We assume that the observed rating consists of two components: (i) a value-for-money rating, Ψ^v, and (ii) an effort rating, Ψ^e.[4] We discuss the generation of ratings below.

Travelers' beliefs, preferences, and demand In each period, there is a continuum of travelers with mass one. Travelers value the apartment's quality, the host's effort during the period of their stay, their idiosyncratic taste for the apartment, and money. The horizontal taste ω_i is uniformly distributed, $\omega_i \sim U[0, 1]$. At the time of purchase, travelers know neither the quality of the apartment nor the effort the host will exert during their stay. However, they have access to the ratings vector $\Psi = \{\Psi^v, \Psi^e\}$.

Based on the ratings Ψ^v and Ψ^e, travelers form a belief about both the apartment's quality and the host's effort. We model the belief formation process in a reduced form. Specifically, let the belief that the apartment is of quality H be $\mu(\Psi^v) \in [0, 1]$ with $\mu(0) = 0$ and $\mu'(\Psi^v) > 0$; that is, higher ratings increase travelers' beliefs about quality. This assumption reflects that travelers associate a higher value-for-money rating—all else equal—with a higher quality apartment. Similarly, we assume that the belief about the host's exerted effort level is $\nu(\Psi^e) \geq 0$ with $\nu(0) = 0$ and $\nu'(\Psi^e) > 0$. Thus, better effort ratings induce travelers to expect more effort to be exerted.

Travelers are risk neutral and have an additively separable utility function resulting in the following expected utility, given their beliefs:

$$u(\Psi^v, \Psi^e, \omega_i, p) = \mu(\Psi^v) + \nu(\Psi^e) + \omega_i - p. \tag{1}$$

We normalize the travelers' outside option to zero. It follows that a traveler will book the apartment if $u(\Psi^v, \Psi^e, \omega_i, p) \geq 0$, which implies that there is an indifferent

[4] It is straightforward to include an overall rating, Ψ^o, into the model setup. For our analysis, we abstract from the overall rating for convenience only. One reasonable assumption would be that Ψ^o is a convex combination of Ψ^v and Ψ^e, which in our setup would render it redundant as long as Ψ^v and Ψ^e are observable. Our implications would be qualitatively similar if, alternatively, travelers only observe an aggregate rating Ψ^o, which is a (known) function $f(\Psi^v, \Psi^e)$.

traveler with cutoff taste $\tilde{\omega} = p - \mu(\Psi^v) - \nu(\Psi^e)$ such that all travelers with $\omega_i \geq \tilde{\omega}$ book the apartment. Given that the taste parameters are uniformly distributed, we obtain the demand function[5]

$$q(\Psi^v, \Psi^e, p) = 1 + \mu(\Psi^v) + \nu(\Psi^e) - p. \tag{2}$$

Note that the host's effort choice does not affect flow revenues in our setup. However, the host's effort choice matters for future profits as it affects the updated ratings that are available to future travelers.

Ratings generation Travelers who book and stay at the apartment generate ratings. We assume that travelers rate non-strategically and, in particular, that every traveler who visits an apartment rates it with a fixed probability independent of any traveler-apartment-specific characteristics.[6] Denoting individual ratings by lower-case letters, i.e., ψ^v and ψ^e, we let the value-for-money rating depend on the apartment's quality, θ, the host's effort in the period of the traveler's stay, e, the traveler's idiosyncratic taste, ω_i, and the price paid for the stay, p. In particular, we assume[7]

$$\psi_i^v = \varphi^v(\theta, e, \omega_i, p) \text{ with } \varphi_\theta^v > 0, \varphi_e^v > 0, \varphi_{\omega_i}^v > 0, \varphi_p^v < 0. \tag{3}$$

Thus, a higher quality, greater effort, and higher taste induce a higher value-for-money rating. In contrast, a higher price induces a lower value-for-money rating.[8] Importantly, by setting a price, p, the host determines the average traveler's taste, ω^e, so that the price has an additional, indirect effect via the selection of travelers who stay at the apartment.

The effort rating is generated by a rating function:

$$\psi_i^e = \varphi^e(e, p) \text{ with } \varphi_e^e > 0, \varphi_p^e \leq 0, \tag{4}$$

so that the effort rating depends positively on the exerted effort and potentially negatively on the price.

For simplicity, we assume that the average rating left by travelers who stay at the apartment is used to update the rating displayed to future travelers. This implies that

[5] We implicitly assume $p \leq 1 + \mu(\Psi^v) + \nu(\Psi^e)$ such that the market is at least partially covered.

[6] While we abstract from any form of selective rating, this would not alter our main effects but instead add an additional effect to our model, similar to that discussed in Carnehl et al. (2022).

[7] To simplify the notation, we denote partial derivatives by using subscripts, i.e., $\frac{\partial f}{\partial x} =: f_x$.

[8] Note that in principle, future travelers could also use the value-for-money rating to form their beliefs about the host's effort. We abstract from this to streamline our discussion.

this average rating is deterministic.[9] Consequently, the average *purchasing* traveler's taste is used to compute the average rating left in any given period. This allows us to write the induced average value-for-money rating as

$$\psi^v(\theta, e, \omega^e, p) = \varphi^v\left(\theta, e, \omega^e(p), p\right), \tag{5}$$

while the effort rating is identical for all travelers, implying $\psi^e(e,p) = \varphi^e(e,p)$.

Ratings are updated from one period to the next as a simple average. Thus, an initial rating of Ψ_t^j in period t is updated to Ψ_{t+1}^j with an incoming (average) rating ψ_t^j according to

$$\Psi_{t+1}^j = \frac{t}{t+1}\Psi_t^j + \frac{1}{t+1}\psi_t^j. \tag{6}$$

We parametrize the rating system to have an initial rating of $\Psi_0 = (\Psi_0^v, \Psi_0^e)$, which induces initial beliefs $\mu(\Psi_0^v)$ an $\nu(\Psi_0^e)$. These can be interpreted as potential travelers' initial attitudes towards the listing based on the apartment's description provided on the platform.

Naive host As a benchmark, we consider a naive host who maximizes flow profits, not considering the effect of price and effort on ratings. Such a host maximizes

$$(1 + \mu(\Psi_t^v) + \nu(\Psi_t^e) - p)p - \frac{c}{2}e^2 \tag{7}$$

with respect to p and e, which yields an optimal price:

$$p_t^m = \frac{1 + \mu(\Psi_t^v) + \nu(\Psi_t^e)}{2}, \quad e_t^m = 0. \tag{8}$$

Thus, a naive host never exerts effort and chooses the monopoly price, given current travelers' beliefs.

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

$$\max_{(p_t,e_t)_{t=1}^T} \sum_{t=1}^{T} \delta^t \left(q(\Psi_t^v, \Psi_t^e, p_t)p_t - \frac{c}{2}e_t^2 \right), \tag{9}$$

[9]The extension to single-unit sales in each period is straightforward. The only noise that would appear derives from the realization of the idiosyncratic component.

subject to the law of motion of the ratings defined above. We rewrite this problem in terms of the corresponding Bellman equation to obtain[10]

$$V_t(\Psi_t^v, \Psi_t^e) = \max_{p_t, e_t} \left(q(\Psi_t^v, \Psi_t^e, p_t)p_t - \frac{c}{2}e_t^2 + \delta V_{t+1}\left(\Psi_{t+1}^v(p_t, e_t), \Psi_{t+1}^e(p_t, e_t)\right)\right). \quad (10)$$

To keep the theoretical part concise, we do not provide a complete characterization of the dynamic problem—in particular, this would require an explicit specification of the belief-updating process. Instead, we illustrate the dynamic incentives arising for the strategic host as this is sufficient to derive testable hypotheses.

Lemma 1 *The effort and value-for-money ratings are both increasing in the host's effort. Moreover, the effort rating is weakly decreasing in the price.*

The impact of the price on the value-for-money rating depends on the relative strength of the value-for-money effect compared to the selection effect. If the value-for-money effect is relatively stronger than the selection effect, a marginal price increase would lead to a decline in the rating:

$$\frac{d}{dp_t}\psi^v < 0 \iff \frac{-\varphi_p^v}{\varphi_\omega^v} > \frac{1}{2}. \quad (11)$$

We relegate all of our derivations to Appendix A. The impact of the host's effort on both ratings and of the price on the effort rating follows immediately from the assumptions on the ratings functions. The condition (11) in turn obtains by noting that a change in the price affects the value-for-money rating both via the direct impact on the induced rating (φ_p^v) and via changing the average traveler's taste ($\varphi_\omega^v \cdot \frac{d\omega^e}{dp} = \frac{1}{2}\varphi_\omega^v$). The relative sizes of the direct and indirect impacts, which go in opposite directions, thus determine the overall impact of a price change on the induced value-for-money rating. If the direct price impact, or value-for-money effect, is relatively low, the selection effect dominates and the induced rating increases in the price. The converse is true if the direct price impact is relatively high.

Lemma 1 naturally has implications for the incentives of the host who dynamically optimizes their rating. To positively influence the future rating stock, the host would have an incentive to exert a strictly positive effort in all but the last period. The pricing incentive, however, depends in particular on the relative strength of the value-for-money and selection effect and the extent to which the price affects the effort rating.

[10] It is straightforward to verify that our setup satisfies the sufficient conditions for this to be feasible.

Lemma 2 *A strategic host exerts a positive effort in every period, $t < T$. Whether a strategic host prices above or below the naive optimal price depends on the sign of*

$$\frac{dV_{t+1}}{d\Psi^v}\left(\varphi_\omega^v \frac{1}{2} + \varphi_p^v\right) + \frac{dV_{t+1}}{d\Psi^e}\varphi_p^e. \tag{12}$$

Equation (12) reflects that the price has an impact not only on the value-for-money rating, where the sign of the impact depends on the sign of $\varphi_\omega^v \frac{1}{2} + \varphi_p^v$, but also on the effort rating. Note that whenever $\varphi_p^e = 0$, that is, when the price *only* affects the value-for-money rating, the sign of (12) depends only on the sign of $\varphi_\omega^v \frac{1}{2} + \varphi_p^v$. As such, the price distortion—relative to the myopically optimal price—depends only on the relative strength of the value-for-money and selection effects. If this were not the case, a price increase would have an additional detrimental impact on the effort rating, which, all else equal, would strengthen the incentive to lower prices in the interests of ratings management.

Prices and effort—substitutes or complements? A naturally arising question is whether and how hosts trade off strategic pricing and effort in their ratings management. In particular, a host can continue to increase their profits via improved ratings by exerting more effort, by adjusting the price according to Lemma 1, or by doing both. We address this question of optimal ratings management by analyzing how the optimal price and effort levels interact strategically. We say that effort and price are strategic complements if an increase in effort induces a further distortion in the price away from the optimal naive price, that is, if an increase in effort increases the incentive to use the price to improve the ratings further.

Lemma 3 *Strategic pricing and effort provision are complements if and only if*

$$sign\left(\frac{d^2V_{t+1}}{de_t dp_t}\right) = sign\left(\frac{dV_{t+1}}{d\Psi^v}\left(\varphi_\omega^v \frac{1}{2} + \varphi_p^v\right) + \frac{dV_{t+1}}{d\Psi^e}\varphi_p^e\right). \tag{13}$$

If they are complements (substitutes), an increase in effort increases (reduces) the incentive to distort the price from the optimal naive price.

The economic forces driving the interaction of strategic pricing and effort are best exemplified by muting the impact of the price on the effort rating along with potential cross effects in the value-for-money rating. Suppose that a host marginally increases their effort. The higher effort improves the value-for-money and effort ratings. The pricing decision depends on the effect the current price has on future ratings. How the host wants to adjust their price in response to a change in their effort depends on the resulting change in the marginal benefit of strategically distorting the price. If this marginal benefit increases, then strategic pricing and effort provision are complements and the seller will further distort the price in the direction that will lead to higher future ratings.

Lemma 3 establishes under which conditions the marginal benefit of ratings management via the price increases (or decreases) in the host's effort. If there were only a single aggregate rating, the condition would correspond to a simple condition on the curvature of the continuation profits—specifically, if the continuation profits are convex (concave) in the aggregate rating, prices and effort are complements (substitutes). The condition is slightly more complex in the general multi-rating setting with potential cross effects, however, the intuition is the same.

For example, suppose that the continuation profits depend on a weighted average of the value-for-money and the effort ratings, $\Psi := \lambda\Psi^v + (1-\lambda)\Psi^e$ with $\lambda \in [0,1]$, and that there are no cross effects in the ratings functions, $\phi^v_{pe} = \phi^e_{pe} = 0$. In this case, Lemma 3 implies that prices and effort are complements (substitutes) if and only if the continuation value is convex (concave) in the aggregate rating.

In that case, an increase in effort leads to better future ratings, increasing the marginal benefit of better ratings when the continuation profits are convex and, thus, amplifying the incentive to use the price to improve future ratings.

Whether strategic pricing and effort are substitutes or complements does not determine whether the model predicts a positive or negative relationship between equilibrium prices and effort levels—this also depends on whether the value-for-money or the selection effect is dominant. For example, strategic pricing and effort being complements would imply a negative relationship between prices and effort—lower prices are associated with increased effort levels—but only if the value-for-money effect dominates the selection effect. In this case, increased effort increases the marginal benefit of engaging in strategic pricing. The latter leads to downward pricing pressure because the dominant value-for-money effect implies that lowering prices increases the ratings obtained in the future.

2.1 Hypotheses for the Empirical Analysis

The theoretical analysis gives rise to several hypotheses which test either the assumptions of the model or its predictions. In formulating these hypotheses, we reflect that the model is deliberately stylized to isolate the key economic forces that are at play. Specifically, we obtain hypotheses about the role of effort and prices and the complementarity between the two.

Hypothesis 1 (Role of Effort) *Both the value-for-money rating and the effort rating are positively affected by the effort.*

Hypothesis 1 essentially comprises the model assumption that the host's effort has a positive impact on both the value-for-money rating and the effort rating.

With respect to the impact of the price on the respective ratings, Lemma 1 offers potentially competing hypotheses. In particular, the model gives scope for prices to either positively or negatively affect the induced value-for-money rating, depending on the relative strength of the direct price and indirect selection effects. While this is ultimately an empirical question that we aim to answer, we conjecture here and in the following hypotheses that the value-for-money effect outweighs the selection effect in the context of Airbnb listings. This conjecture is in line with recent work by Luca and Reshef (2021), who find an analogous correlation in the context of restaurant ratings found on Yelp.

Hypothesis 2 (Role of Price) *Both the value-for-money rating and the effort rating are negatively affected by the price.*

With respect to Hypothesis 2, it is of particular interest to see whether this conjectured negative relationship obtains for all types of listings, or whether listings of particular characteristics (such as high-value or high-price listings) exhibit differential behaviors. Within our model, this would be in line with a relatively stronger product-specific selection effect for these types of listings or a weaker price sensitivity in the rating behaviors of travelers selecting into booking such apartments.

Similar to the impact of the price on the value-for-money rating, Lemma 3 offers potentially competing hypotheses regarding the complementarity or substitutability of strategic pricing and effort. This is driven primarily by the curvature of the continuation profits at particular rating levels. We consider it plausible that the marginal benefit of higher ratings is initially high but that it decreases for sufficiently high ratings and increases otherwise. This would imply that strategic pricing and effort provision are complements for listings with low ratings but become substitutes once the listings have achieved a sufficiently high rating. As discussed, the sign of the correlation between price and effort, given the complementarity or substitutability, depends on whether the value-for-money or selection effect dominates.

Hypothesis 3 *Strategic pricing and effort are complements for low values of the value-for-money rating and substitutes for high values of said rating.*

If Hypothesis 2 holds, Hypothesis 3 predicts that price and effort measures are negatively correlated for low values of the value-for-money rating and positively correlated for high values of the value-for-money rating.

Finally, the model offers predictions regarding the dynamic behavior of hosts. In line with Hypothesis 2, we condense the potentially ambiguous model predictions by conjecturing that strategic hosts overall charge lower prices than naive sellers because the value-for-money effect dominates.

Hypothesis 4 *Strategic hosts charge lower prices than naive hosts, conditional on a given rating stock. However, strategic hosts maintain a higher rating than naive hosts, allowing them to charge higher prices after a sufficient number of bookings.*

The first part of this hypothesis follows directly from Hypotheses 1 and 2. The second part of Hypothesis 4 is a consequence of the ratings aggregation mechanism: over time, the impact of an additional incoming rating decreases.

Overall, we expect that strategic hosts charge lower prices within-period, compared to naive hosts, to positively affect their rating stock. Such strategic behavior implies that as their ratings accumulate, the strategic hosts' ratings will reach higher values than those of the naive hosts. But as the ratings become sufficiently unresponsive over time, the dynamic pricing incentives become weaker and the direct effect that higher ratings allow for higher prices dominates the strategic pricing incentives.

3 The Data

To test the assumptions and predictions of our model, we combine transactions and ratings data on Airbnb listings in Paris, France. Our observations span the entire year 2017.

The transactions data we use for our study are obtained from AirDNA, a specialist for short-term rental data.[11] 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.

The ratings data are obtained from InsideAirbnb.com.[12] 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

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

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

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.

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.[13]

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 ratings. 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.[14] Importantly, leaving a rating is optional so not all stays are rated. To discard bookings that did not receive a rating, we use timestamp information revealing when ratings were submitted. Guests can leave a rating on Airbnb within 14 days of

[13]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.

[14]The rating information contained in the Inside Airbnb data is provided at the beginning of each month.

their stay. Therefore, if a new rating appears within 14 days of a booking, we label this booking as "rated." If there are multiple bookings in the 14 days prior to the rating, we choose the closest one to the rating date. When studying the impact of prices on ratings, we calculate the monthly average price using only rated 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. 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 rated.[15] We remove the upper percentile of the price distribution and normalize all prices by the number of guests included in the price.[16] 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. 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 $r \in [95, 100]$.

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 ratings, which is in line with anecdotal evidence.[17] This implies

[15] A study by Fradkin et al. (2021) reports that 68 percent of bookings receive ratings. However, the geographical scope of their data is not specified. The ratio of ratings 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).

[16] The InsideAirbnb data provide information on how many travelers are included in the price a host posts on the Airbnb website. The data contain some extreme price outliers, which we remove by trimming the data at the upper percentile.

[17] See hhttps://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 ratings. We provide supporting empirical evidence in Appendix B.

that any effect of prices on ratings should be less pronounced for listings that received more ratings because the marginal impact of one individual rating on the average rating diminishes as more ratings accumulate.

4 Empirical Analysis

In the following empirical analysis, we test the hypotheses outlined in Section 2. Our model predictions and the derived hypotheses crucially depend on the overall impact of prices on ratings. Therefore, we first provide evidence that the value-for-money effect dominates the selection effect for most rating categories.[18]

4.1 Value-for-Money vs. Selection Effect

Figure 1 shows the cumulative distribution of the overall star rating for listings in different price terciles. These price terciles were computed based on the average prices observed for each listing. Figure 1 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, less than 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 only 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 ratings on period $t-1$ prices. We argue that, after taking into account these two potential explanations, the remaining conditional

[18] 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.

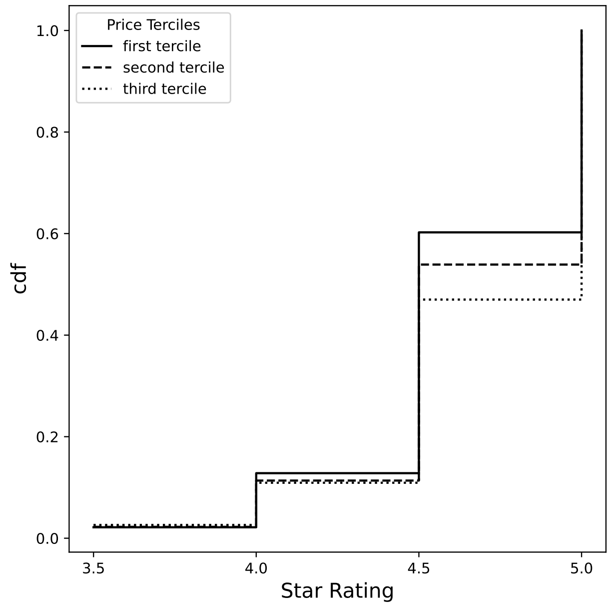


Figure 1: Rating Distribution for Different Price Terciles

.

Notes: Price terciles are computed based on average listing prices. For exposition, we only show the cumulative distribution functions (CDFs) starting from a rating of 3.5 stars.

correlation between ratings and prices is indicative of the relative importance of the value-for-money and the selection effect.

Specifically, we estimate the following equation:

$$r_{it}^{cat} = \beta_0 + \beta_1 \times log(p_{it-1}) + X_{it}'\gamma + \mu_i + \mu_t + \epsilon_{it}, \tag{14}$$

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 occur. This time lag between prices and ratings helps address reverse causality concerns. To test Hypothesis 2, 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, 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. Its coefficient is of interest to test Hypothesis 1. 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}.

Table 2: Fixed-Effects Regression of Ratings on Prices

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
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)
log(resp. rate)	0.017^{**}	0.012	0.008	0.014^{*}	0.012	0.015^{*}	0.006
	(0.006)	(0.008)	(0.007)	(0.008)	(0.009)	(0.007)	(0.007)
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 the listing fixed effects, month fixed effects, and a second-order polynomial of the number of ratings. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively.

Table 2 shows the main coefficients of interest from estimating Equation (14) 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 the least correlated with prices. Higher effort correlates positively with ratings. Crucially, the price negatively correlates with the overall rating, which is the most salient rating to customers. This suggests that the value-for-money effect indeed dominates the selection effect in our data.

Table 3 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 3 are consistent with a higher price sensitivity in the ratings of price-conscious customers, which, according to our model, should reinforce the value-for-money effect of prices on ratings.[19]

The results of Tables 2 and 3 are consistent with a dominant value-for-money effect according to which sophisticated hosts can increase their ratings through lower prices. The incentive to lower prices depends on the trade-off between foregone revenues from lower prices today and increased revenues from higher demand in the

[19] We present robustness checks based on a first-differences estimator for the results of Tables 2 and 3 in Appendix D.

Table 3: Fixed-Effects Regression of Ratings on Low and High Prices

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
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)
log(resp. rate)	0.017^{**}	0.012	0.008	0.014	0.012	0.015^{*}	0.006
	(0.006)	(0.008)	(0.007)	(0.008)	(0.009)	(0.007)	(0.007)
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. Note that the price coefficients show the slope for each price category. Standard errors are shown in parentheses and are clustered at the listing level. All regressions further include listing fixed effects, month fixed effects, and a second-order polynomial of the number of ratings. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication ratings, respectively.

future. All else equal, receiving a reward in the more distant future mitigates the incentive to cut prices immediately compared to receiving the same reward sooner. Thus, hosts closer to reaching a threshold resulting in a salient overall rating increase should have a stronger incentive to exploit the value-for-money effect and to charge relatively lower prices.

To analyze whether our data support this hypothesis, we estimate the following equation:

$$p_{it} = \beta_0 + \sum_{j=lb}^{ub} \beta_j \mathcal{I}(r_{it} = j) + q_{it} + l_i + \mu_t + \epsilon_{it}, \tag{15}$$

where p_{it} denotes the offer price of listing i in period t, r_{it} denotes the granular overall rating of listing i at the *beginning* of period t,[20] q_{it} 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_{it} are obtained from a hedonic regression of offer prices on listing fixed effects, and interaction terms between deciles for the number of ratings and an indicator variable for five-star ratings. Based on the estimated fixed effects

[20]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.

and the coefficients of the interaction terms, we obtain the quality indicators by computing the quintiles of the distribution of the predicted prices.[21]

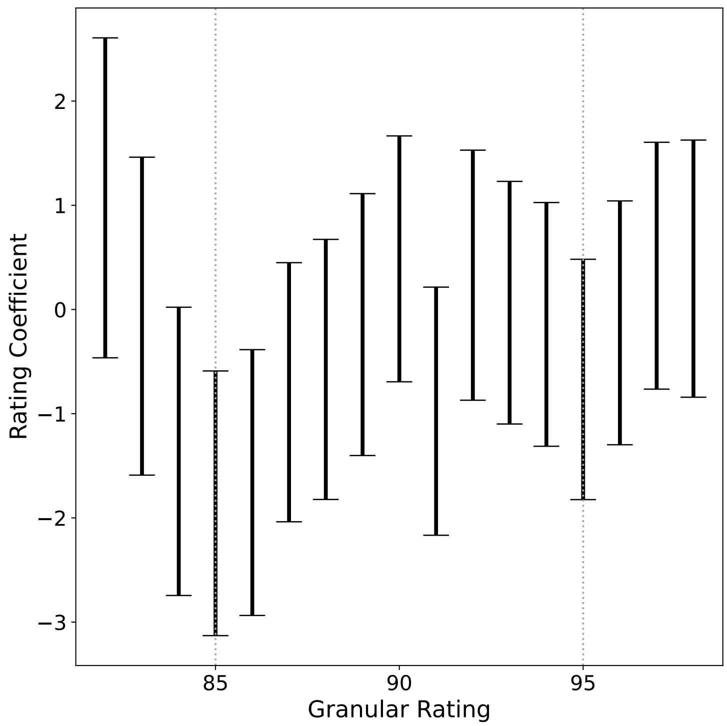


Figure 2: Conditional Average Offer Prices Around Salient Threshold

Notes: The error bars show the 95% confidence intervals using heteroskedasticity-robust standard errors.

Figure 2 shows the rating coefficients we obtain when estimating Equation (15) for $j \in [82, 98]$ (using the granular rating of 81 as the baseline category). 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. The estimation relies on listings for which the number of yearly bookings is larger than the median. The motivation for this choice is that more-active hosts are more likely to engage in strategic behavior than less-active, occasional hosts.

Figure 2 reveals a V-shaped pricing pattern around the salient ratings thresh-

[21]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.

olds.[22] The monotonic price decreases observed before each threshold are consistent with the mechanism outlined in Hypothesis 4. 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 lose their newly obtained higher salient ratings due to the dominant value-for-money effect.[23] This is in line with findings in concurrent work by Zhong (2022), who similarly finds a V-shaped pricing pattern around ratings thresholds on the online shopping platform Taobao.

We conclude this section by noting that our results do not preclude the presence of a selection effect in the generation of ratings. However, our results suggest that the value-for-money effect dominates the selection effect. Based on our theoretical framework, this insight allows us to test the hypotheses on the price-effort relationship, as well as dynamic price and ratings patterns, as it determines the hosts' incentives and thus the predicted sign of the relationships between the variables of interest.

4.2 Price-Effort Relationship

The results in Section 4.1 suggest that the value-for-money effect dominates the selection effect. Consequently, sellers can obtain higher ratings and hence higher future profits by strategically lowering the price of a listing at the expense of lower flow profits. This insight is important when testing Hypothesis 3. The hypothesis posits that strategic pricing and effort are complements for lower levels of the value-for-money rating and substitutes for higher levels of this rating. This hypothesis builds on the assumption that the continuation value is convex in ratings for low rating levels and becomes concave for high levels.

To test the hypothesis, we proceed in two steps. We first provide evidence that the continuation profit as a function of the overall rating displays a convex-concave relationship. In a second step, we explicitly analyze the relationship between prices and effort. Our model predicts a negative (positive) relationship between prices and effort whenever (i) the continuation profits are convex (concave) in the rating and (ii) the value-for-money effect dominates the selection effect. Thus, if the continuation profits exhibit a convex-concave shape, we expect a negative correlation to arise for low rating levels and a positive correlation for high rating levels.

[22] The V-shaped pattern observed in the vicinity of the thresholds does not depend on the overall range chosen to estimate Equation (15).

[23] We provide additional robustness checks for the results shown in Figure 2 in Appendix E.

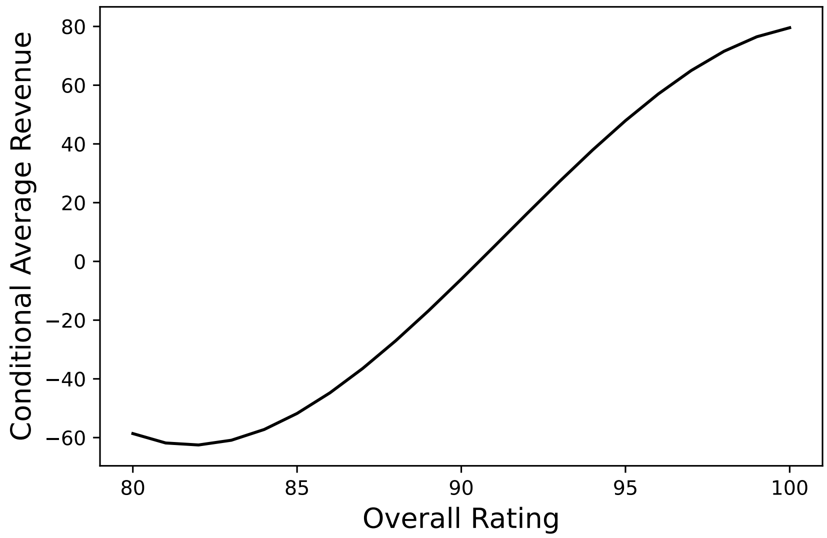


Figure 3: Relationship Between Monthly Revenues and Overall Ratings.

Notes: Only listings with a granular rating above 80 are used for the estimation. The curve is obtained from the parameters of the third-degree polynomial and by adjusting the mean to zero.

To gain insights into the relationship between the continuation profits and the ratings, we regress the total monthly revenue (pq_t) of listings on a third-degree polynomial of the granular rating measure at the *beginning* of the month (r_t), listing fixed effects (μ_i), month fixed effects (μ_t), and time-varying controls (X_{it}). These controls include the number of days a listing is available during a given month, the logarithm of the host's response rate, and the deciles of the distribution of the number of ratings.

$$pq_{it} = \beta_0 + \beta_1 \times r_{it} + \beta_2 \times (r_{it})^2 + \beta_3 \times (r_{it})^3 + X'_{it}\gamma + \mu_i + \mu_t + \epsilon_{it}\,. \tag{16}$$

Since the continuation profits are the discounted sum of monthly profits, we view the parameters of the third-degree polynomial in Equation (16) as a reasonable approximation of the relationship between the granular ratings measure and the continuation value.[24] The estimated relationship between ratings and revenues based

[24]Note that at least a third-degree polynomial is necessary to allow for the possibility of convexity and concavity over the domain of a non-decreasing function. The choice of a third-order polynomial in Equation (16) can be justified when analyzing the relationship non-parametrically. We provide additional details on the non-parametric estimation in Appendix F.

Table 4: Price-Effort Regression Results

	Comm.	Clean.	Check-in	Resp. rate
log(price)	-0.041^{***}	-0.001	-0.027^{*}	-0.835^{*}
	(0.011)	(0.014)	(0.012)	(0.347)
log(price) $\times$ $\mathcal{I}_{q=2}$	0.026^{**}	0.026^{*}	0.014	0.280
	(0.011)	(0.013)	(0.011)	(0.221)
log(price) $\times$ $\mathcal{I}_{q=3}$	0.046^{***}	0.019	0.029^{*}	-0.293
	(0.013)	(0.015)	(0.013)	(0.267)
log(price) $\times$ $\mathcal{I}_{q=4}$	0.049^{***}	0.020	0.029^{*}	-0.065
	(0.014)	(0.017)	(0.015)	(0.301)
log(price) $\times$ $\mathcal{I}_{q=5}$	0.078^{***}	-0.006	0.038^{*}	-0.170
	(0.016)	(0.019)	(0.018)	(0.327)
Obs.	337,005	337,236	336,539	320,103

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 further include listing fixed effects, month fixed effects, a second-order polynomial of the number of ratings, and constants for the respective quintiles. Comm., Clean., and Resp. rate denote the communication rating, the cleanliness rating, and the response rate, respectively.

on the estimated parameters of the ratings polynomial is shown in Figure 3. The results corroborate a convex relationship in the lower range of the domain and a concave relationship in the upper range.

Based on these empirical findings, the model predicts a strategic complementarity between prices and effort in the lower range of the ratings distribution and strategic substitutability in the upper range. As previously explained, the dominant value-for-money effect implies a negative correlation between price and effort measures for lower levels of the ratings distribution and a positive correlation in the upper range.

To assess the model prediction, we create five bins defined by the quintiles of the granular ratings distribution. Define as q_k the kth ratings bin defined by the respective quintiles of the granular ratings distribution.[25] We estimate the following regression:

[25] We use quintiles of the granular ratings, rather than the aggregate the star ratings, to have more variation in the data. Recall that the median star rating is 4.5.

$$e_{it} = \alpha_1 + \sum_{k=2}^{k=5} \alpha_k \mathcal{I}\{r_i \in q_k\} + \beta_1 log(p_{it-1}) + \sum_{k=2}^{k=5} \beta_k log(p_{it-1}) \mathcal{I}\{r_i \in q_k\} + X'_{it}\gamma + \mu_i + \mu_t + \epsilon_{it}\,, \tag{17}$$

where r_{it} denotes the granular overall rating of listing i at time t, $\mathcal{I}$ is the indicator function which takes the value one if the rating lies in the respective bin defined by the quintiles, and p_{it-1} denotes the offer price in period $t-1$. As usual, we control for listing and month fixed effects. As proxies for the effort level e chosen by the host, we use the three rating categories associated with effort (cleanliness, communication, and check-in) and the host's response rate. The X matrix contains the second-order polynomial of the number of ratings.

The results are displayed in Table 4. We observe a pattern consistent with decreasing complementarities for the communication and check-in ratings. For the lowest rating quintile, prices and effort are negatively related, while their association becomes increasingly positive for higher quintiles. While we observe the same pattern for the response rate in the lowest quintile, it does not hold overall. However, the response rate captures only a specific aspect of the overall effort hosts dedicate to their guests and, in particular, might be related to situations in which the host knows that no transaction will take place.[26] Effort ratings might provide a more complete measure by also incorporating aspects not captured by direct measures, such as friendliness or level of detail of information.

As an alternative way of looking at the host's strategic usage of prices and effort, we look at how effort and value-for-money ratings interact. Table 5 shows the conditional correlation between the effort-related ratings and the value-for-money rating for the different quintiles of the overall rating. The complementarity between prices and effort suggests that the value-for-money rating and the respective effort ratings should be positively correlated for lower quintiles of the overall rating. As effort and prices become increasingly substitutable, the positive correlation should diminish. That is, once the ratings have become sufficiently positive, the host can slack in one of the strategic ratings-management instruments while exercising discipline regarding the other. The results from table Table 4 are, overall, in line with this intuition.

[26]For example, when the host receives messages requesting a stay for a particular date although the listing is already booked.

Table 5: Effort-Value Regression Results

	Comm.	Clean.	Check-in	Resp. rate
Value	0.189***	0.272***	0.198***	−0.317
	(0.010)	(0.012)	(0.011)	(0.170)
Value × $\mathcal{I}_{q=2}$	−0.52***	−0.085***	−0.075***	0.211
	(0.011)	(0.013)	(0.012)	(0.221)
Value × $\mathcal{I}_{q=3}$	−0.115***	−0.128***	−0.118***	0.175
	(0.011)	(0.013)	(0.012)	(0.223)
Value × $\mathcal{I}_{q=4}$	−0.144***	−0.160***	−0.143***	0.120
	(0.011)	(0.014)	(0.013)	(0.260)
Value × $\mathcal{I}_{q=5}$	−0.088***	−0.123***	−0.099***	0.032
	(0.015)	(0.017)	(0.016)	(0.259)
Obs.	336,425	336,374	336,380	319,165

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 further include listing fixed effects, month fixed effects, a second-order polynomial of the number of ratings, and constants for the respective quintiles. Comm., Clean., and Resp. rate denote the communication rating, the cleanliness rating, and the response rate, respectively.

5 Price-Rating Dynamics

In this section, we turn to a dynamic analysis of the relationship between prices and ratings. Hypothesis 4 proposes that strategic hosts offer relative discounts when entering the platform. In light of the dominant value-for-money effect established by our previous analyses, hosts charging low entry prices should 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.

Unfortunately, we do not have a clear explicit measure to differentiate between strategic and naive hosts. To address this issue, we analyze whether listings that charge a relatively lower price when entering the platform can benefit from this in later periods. By establishing that this is the case, we provide evidence consistent with Hypothesis 4 in that entry price discounts are profitable to hosts, with strategic hosts being more likely to adopt such a profitable strategy.[27]

To analyze the impact of entry price discounts, we estimate the following equation:

$$y_{it} = \sum_{\tau=1}^{\tau=6} \alpha_\tau \mathcal{I}(t=\tau) + \sum_{\tau=1}^{\tau=6} \beta_\tau \times fp_i \times \mathcal{I}(t=\tau) + X_{it}'\gamma + \epsilon_{it}\,, \tag{18}$$

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. fp_i measures the initial discount listing i offers in the first period it enters the market.[28] X_{it} contains month and location fixed effects, a fixed effect for the first month in which a listing was observed and, when the outcome variable is the number of bookings or total revenue, the total number of days a listing was available and the number of ratings at the beginning of the respective month.[29] 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 G, we show the results obtained when including listing fixed effects in Equation (19). The fixed effect analysis is consistent with the findings presented here, mitigating concerns related to potential endogeneity.

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

$$fp_i = 1 - \frac{p_{i1}}{\bar{p}_{i2-6}}\,, \tag{19}$$

where $\bar{p}_{i2-6}$ denotes the average price charged after the first month. A negative value of fp_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. β_τ in Equation (18) can be interpreted as the

[27]In Appendix G, we show that hosts offering entry discounts are more likely to acquire the *superhost* status and to have the "instant bookable" option activated. Instant booking allows guests to book an Airbnb accommodation similar to how they would book a hotel room; that is, the booking is immediately confirmed and binding and does not require a review and manual confirmation by the Airbnb host. These results suggest that it is indeed hosts who are more professional—and therefore more likely to act strategically—who are more likely to offer initial discounts.

[28]The idea of interacting a cross-sectionally varying variable with the period fixed effects is inspired by Huber et al. (2021), who apply this approach in the context of discrimination in Nazi Germany.

[29]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.

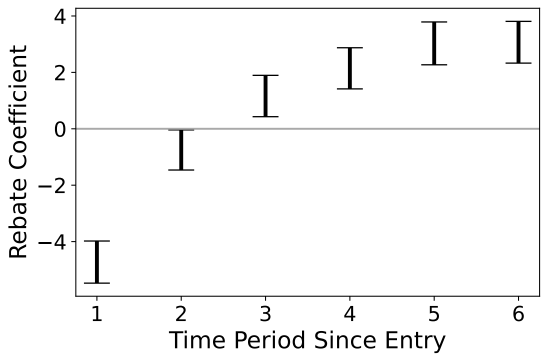
effect of a 100 percent first-period discount on the outcome in period t. To aid the interpretation of the following results, we report the estimates scaled by the median discount (7 percent).

To restrict the sample to new entrants, we only use listings that do not have more than six ratings when we first observe them in our sample. To have a sufficient number of observations to study changes over time, we only use those listings we observe for at least six months, with bookings occurring in each of these six months. Finally, we only include the first six monthly observations for each listing to have a balanced panel.

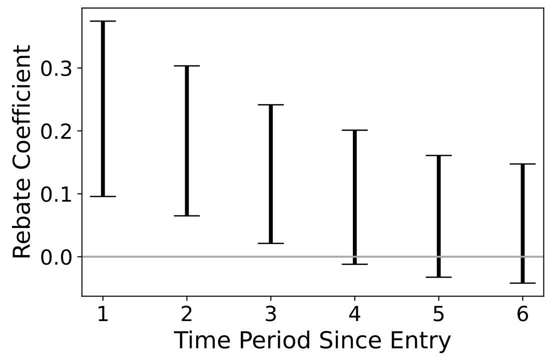
Figure 4 shows the estimated β_t-coefficients multiplied by the median discount for the different outcome variables. Figure 4a shows that the first-period discount of approximately five euros vanishes in the second period and subsequently reverses. While the construction of the discount variable partially predetermines the shape of the curve, this is not the case for its level. For example, if the curve were to lie strictly below zero, it would indicate that listings which generally charge lower prices are more likely to offer an initial discount.

Figure 4b shows that an initial discount is correlated with a higher overall rating. In the first months, those listings that offer an initial discount receive, on average, better overall ratings. This is consistent with the dominant value-for-money effect documented in Section 4.1 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 a persistent boost in the rating score, which dissipates only gradually. The pattern looks broadly similar—albeit estimated with lower statistical precision—for most other rating categories, except for the location rating, see Figure 4f. This is intuitive, as we expect the selection effect to play a larger role in the location rating. We report the results for other rating categories in Appendix G.

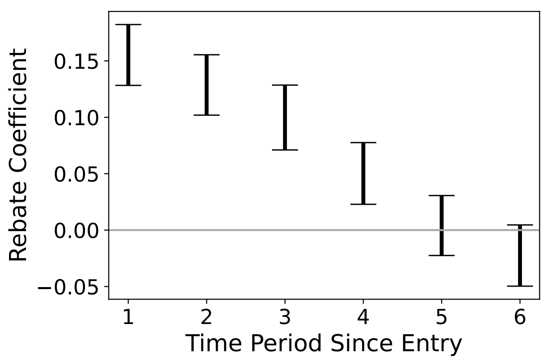
Consistent with a downward sloping demand curve, Figure 4c shows that the initial discount seems to draw additional bookings in the first periods. However, even though the prices of listings with initial discounts are relatively higher in later periods, they do not experience lower numbers of bookings. 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 (see Figure 13b and



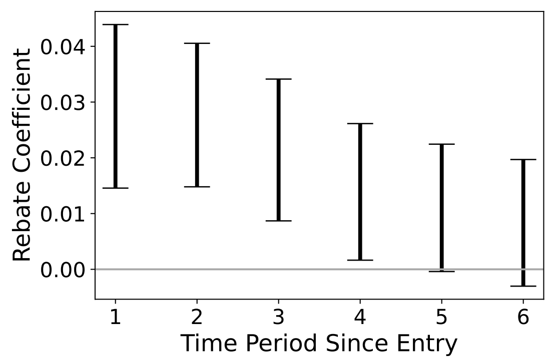
(a) y_{it} = Booking Price



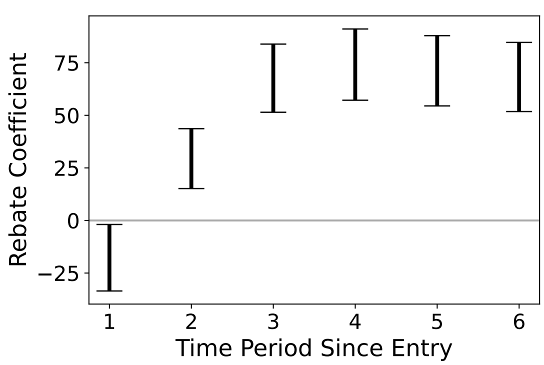
(b) y_{it} = Overall Rating



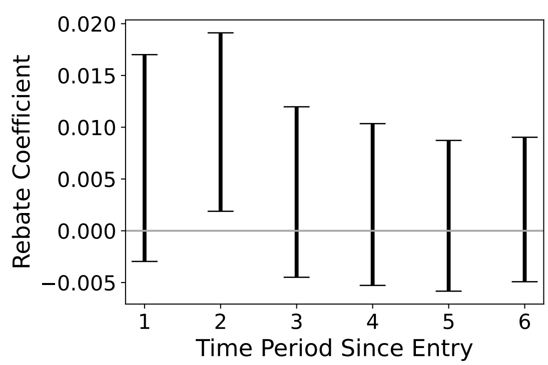
(c) y_{it} = Number of Bookings



(d) y_{it} = Value-for-Money rating



(e) y_{it} = Revenue



(f) y_{it} = Location Rating

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

Notes: The error bars show the 95% confidence intervals using standard errors clustered at the listing level.

Figure 13a for an analysis using the natural logarithms of the number of bookings and the revenue as dependent variables, respectively).

6 Conclusion

Our paper provides a framework that reconciles prior results on the relationship between prices and ratings. We investigate whether hosts on the short-term accommodation platform Airbnb can influence their ratings through strategic price setting. In a stylized theoretical framework, we illustrate two opposing effects of a higher price on ratings. First, higher prices directly result in lower ratings due to a reduced value for money. Second, higher prices indirectly result in higher ratings 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 ratings 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 ratings 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 an arguably stronger selection effect for the location rating, which derives from travelers' idiosyncratic preferences for specific locations in the city. These results suggest that hosts can strategically reduce prices to improve their future ratings.

Our model further posits that hosts can use effort as another strategic control variable to affect their ratings. Whether strategic pricing and effort provision are strategic complements or substitutes depends on whether future profits are convex or concave in ratings.

We find that continuation profits are convex for lower ratings and become concave for high ratings. Given this result, our theoretical model predicts that low-rated hosts should use effort as a strategic complement to pricing. In contrast, high-rated hosts should use effort as a strategic substitute for pricing. We find empirical support for this prediction.

Given the dominant value-for-money effect, we expect that hosts can benefit from strategically offering discounts upon entry. Our analysis of entry pricing confirms this expectation. Listings with relatively lower entry prices receive better value-for-money ratings and more bookings early on, allowing hosts to charge higher prices and realize higher revenues in subsequent periods. Overall, a median discount of seven percent in the month of entry leads to three percent higher revenues six months down the line.

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

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. (2022). Pricing for the Stars: Dynamic Pricing in the Presence of Rating Systems. 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*.

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. (2021). Impact of pricing on online reviews: A field experiment. *Working Paper*.

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.

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).

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 Derivations

A.1 Strategic Seller Problem

Consider the first-order conditions from the Bellman equation

$$1 + \mu(\Psi_t^v) + \nu(\Psi_t^e) - 2p_t \tag{20}$$

$$+ \delta \left(\frac{d}{d\Psi^v} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^v}{dp_t} + \frac{d}{d\Psi^e} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^e}{dp_t} \right) = 0 \tag{21}$$

$$-ce_t + \delta \left(\frac{d}{d\Psi^v} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^v}{de_t} + \frac{d}{d\Psi^e} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^e}{de_t} \right) = 0. \tag{22}$$

Clearly, the value is increasing in the state variables—the ratings Ψ_t^v and Ψ_t^e—as the flow profits and future states Ψ_{t+1}^v and Ψ_{t+1}^e are strictly increasing in both. Thus, it follows immediately that the dynamic incentives are determined by the reaction of the ratings to the control variables price and effort. Note that it follows from our specification of the ratings generation that

$$\frac{d}{dp_t} \Psi_{t+1}^j = \frac{1}{t+1} \frac{d}{dp_t} \psi^j \tag{23}$$

$$\frac{d}{de_t} \Psi_{t+1}^j = \frac{1}{t+1} \frac{d}{de_t} \psi^j. \tag{24}$$

The per-period value for the money rating is affected by the price according to[30]

$$\begin{aligned} \frac{d}{dp_t} \psi^v &= \frac{d}{dp_t} \varphi^v(\theta, e, \omega^e(p), p) \\ &= \varphi_\omega^v \frac{d\omega^e}{dp_t} + \varphi_p^v \\ &= \varphi_\omega^v \frac{1}{2} + \varphi_p^v \end{aligned} \tag{26}$$

[30]Note that the average purchasing traveler has the idiosyncratic taste component

$$\begin{aligned} \omega^e &= \frac{1 + \tilde{\omega}(\Psi^v, \Psi^e, p)}{2} \\ &= \frac{1 + p - \mu(\Psi^v) - \nu(\Psi^e)}{2}. \end{aligned} \tag{25}$$

implying that the overall effect of the price on the listing's value-for-money rating can be either positive or negative. The effect is negative whenever $\frac{-\varphi_p^v}{\varphi_\omega^v} > \frac{1}{2}$, that is, whenever the direct effect of the price is sufficiently large relative to the effect of the taste component.

The per-period value-for-money rating is affected by the host's effort according to

$$\frac{d}{de_t}\psi^v = \frac{d}{de_t}\varphi^v(\theta, e, \omega^e(p), p) = \varphi_e^v > 0, \tag{27}$$

implying that a higher effort always leads to a higher value-for-money rating. The effort rating in turn is affected by the price and effort according to

$$\frac{d}{dp_t}\psi^e = \frac{d}{dp_t}\varphi^e(\theta, e, \omega^e(p), p) = \varphi_p^e \leq 0 \tag{28}$$

$$\frac{d}{de_t}\psi^e = \frac{d}{de_t}\varphi^e(\theta, e, \omega^e(p), p) = \varphi_e^e > 0. \tag{29}$$

We can rewrite the first-order conditions as

$$1 + \mu(\Psi_t^v) + \nu(\Psi_t^e) - 2p_t + \frac{\delta}{t+1}\left(\frac{d}{d\Psi^v}V_{t+1}\left(\varphi_\omega^v\frac{1}{2} + \varphi_p^v\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi_p^e\right) = 0 \tag{30}$$

$$-ce_t + \frac{\delta}{t+1}\left(\frac{d}{d\Psi^v}V_{t+1}\varphi_e^v + \frac{d}{d\Psi^e}V_{t+1}\varphi_e^e\right) = 0. \tag{31}$$

It follows immediately that the host has an incentive to exert effort due to the rating system as

$$e_t = \frac{\delta}{c(t+1)}\left(\frac{d}{d\Psi^v}V_{t+1}\varphi_e^v + \frac{d}{d\Psi^e}V_{t+1}\varphi_e^e\right) > 0. \tag{32}$$

The effect on the price, however, is ambiguous. We obtain

$$p_t = \underbrace{\frac{1 + \mu(\Psi_t^v) + \nu(\Psi_t^e)}{2}}_{p_t^m} + \frac{\delta}{2(t+1)}\left(\frac{d}{d\Psi^v}V_{t+1}\left(\varphi_\omega^v\frac{1}{2} + \varphi_p^v\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi_p^e\right). \tag{33}$$

Inspecting (33), observe that $\frac{\delta}{c(t+1)} > 0$ and hence that whether a strategic host chooses higher or lower prices than a myopic host depends on the sign of

$$\frac{d}{d\Psi^v}V_{t+1}\left(\varphi_\omega^v\frac{1}{2} + \varphi_p^v\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi_p^e. \tag{34}$$

In particular, prices are lower than the myopic prices whenever

$$-\frac{\varphi_\omega^v \frac{1}{2} + \varphi_p^v}{\varphi_p^e} > \frac{\frac{d}{d\Psi^e} V_{t+1}}{\frac{d}{d\Psi^v} V_{t+1}}. \tag{35}$$

Moreover, for the special case where the price does not affect the effort rating, that is, when $\varphi_p^e = 0$, the sign of (34) is fully determined by the sign of $\varphi_\omega^v \frac{1}{2} + \varphi_p^v$, so that

$$p_t < p_t^m \iff \frac{-\varphi_p^v}{\varphi_\omega^v} > \frac{1}{2}. \tag{36}$$

This condition is the same as in (26), which is intuitive—as the price only affects the induced average value-for-money rating, ψ^v, the optimal strategic price is lower than the myopically optimal price if and only if the induced rating, ψ^v, is negatively affected by the price, that is, if and only if the direct price effect dominates the indirect selection effect.

A.2 Strategic Pricing and Effort

To analyze this relationship, consider the effect that a change in effort has on the change in the optimal price, which we can derive from the first-order conditions (30) and (31).

$$\begin{aligned}
\frac{dp_t}{de_t} &= \frac{\delta}{2} \frac{d}{de_t} \frac{dV_{t+1}}{dp_t} \\
&= \frac{\delta}{2} \frac{d^2 V_{t+1}}{de_t dp_t} && (37) \\
&= \frac{\delta}{2(t+1)^2} && (38) \\
&\left(\frac{d^2 V_{t+1}}{d(\Psi^v)^2} \frac{d\psi^v}{dp_t} \frac{d\psi^v}{de_t} + \frac{d^2 V_{t+1}}{d(\Psi^e)^2} \frac{d\psi^e}{dp_t} \frac{d\psi^e}{de_t} + 2 \frac{d^2 V_{t+1}}{d\Psi^v d\Psi^e} \left(\frac{d\psi^e}{dp_t} \frac{d\psi^v}{de_t} + \frac{d\psi^e}{de_t} \frac{d\psi^v}{dp_t} \right) \right) \\
&+ \frac{\delta}{2(t+1)^2} \left(\frac{dV_{t+1}}{d\Psi^v} \frac{d^2 \psi^v}{dp_t de_t} + \frac{dV_{t+1}}{d\Psi^e} \frac{d^2 \psi^e}{dp_t de_t} \right). && (39)
\end{aligned}$$

Whether effort and strategic price adjustments are substitutes or complements depends both on the sign of (39) and the sign of (34)—we say that strategic price management and effort are complements if an increase in effort leads to a further price adjustment in the direction that would increase future profits at the expense of flow profits.

To better understand the economic forces at play, suppose that the continuation profits depend only on a weighted average of the ratings, $V(\Psi^v, \Psi^e) = V(\Psi)$ with $\Psi := \lambda\Psi^v + (1-\lambda)\Psi^e$ and $\lambda \in [0, 1]$. Moreover, suppose that there are no cross-effects between prices and effort in the rating function, that is, that $\varphi^e_{pe} = 0$ and $\varphi^v_{pe} = 0$. This allows us to write

$$\frac{dp_t}{de_t} \propto \frac{d^2 V_{t+1}}{d\Psi^2} \left(\lambda \frac{d\Psi^v}{dp} + (1-\lambda)\frac{d\Psi^e}{dp}\right) \left(\lambda \frac{d\Psi^v}{de} + (1-\lambda)\frac{d\Psi^e}{de}\right). \tag{40}$$

The last two terms in parentheses are the marginal effects of the price and effort, respectively, on future ratings. Thus, whenever the continuation profits are (locally) convex, a marginal increase in effort induces the host to adjust the price further in the direction that leads to higher ratings, that is, prices and effort are complements. The reason is that the convexity of the continuation profits implies that the marginal benefit from higher ratings is increasing in the rating level. As higher effort levels raise the resulting ratings, the incentive to boost ratings further with a rating-improving price adjustment is amplified. If the continuation profit were concave, an effort increase would mitigate this incentive as the marginal benefit to high ratings would be decreasing.

B Aggregation of Individual Ratings

Figure 5 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.

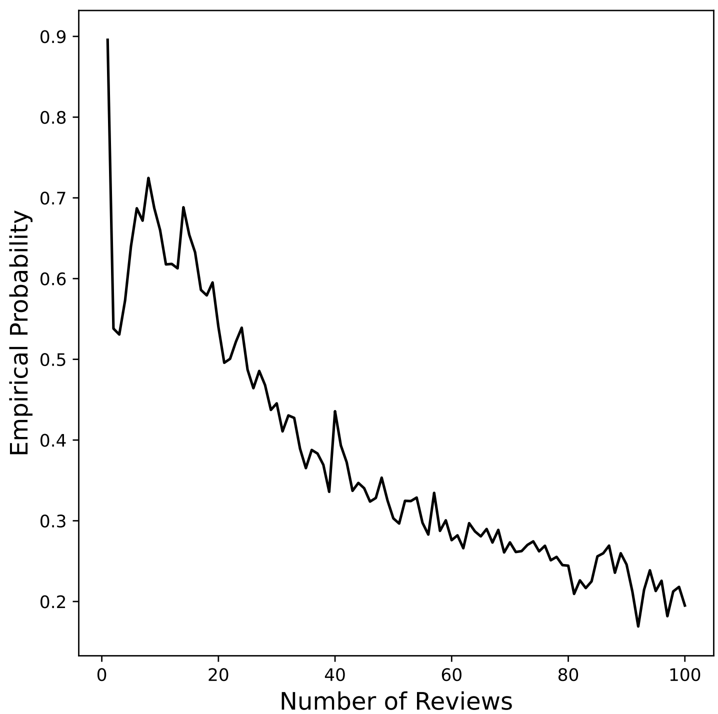


Figure 5: Empirical Probability of Observing a Rating Change

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:

$$y_{it} = \beta_0 + \beta_1 \mathcal{I}_{bw}(r_{it} > \tau) + X_{it}'\gamma + \mu_i + \mu_t + \epsilon_{it}, \tag{41}$$

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.

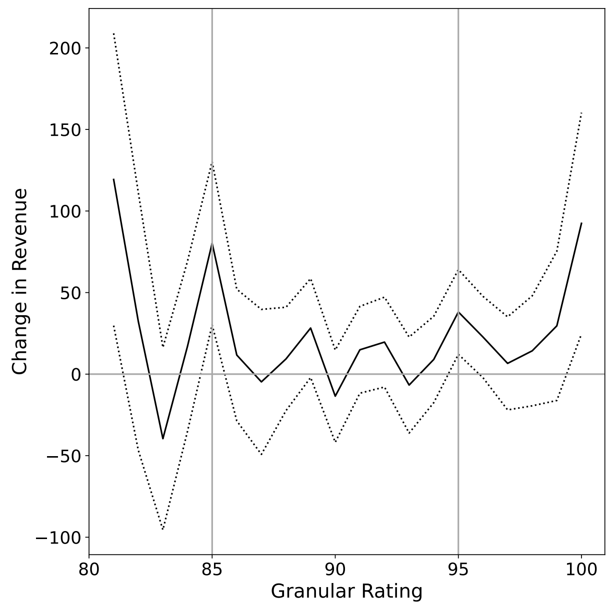


Figure 6: Impact of Incremental Granular Rating Change Using All Listings

Notes: The confidence bands show the 95% confidence intervals using heteroskedasticity-robust standard errors.

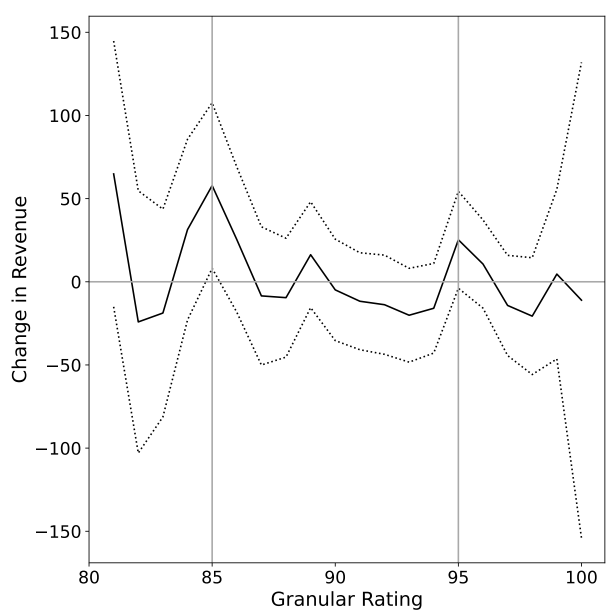


Figure 7: Impact of Incremental Granular Rating Change Using Only Listings that Experience Rating Changes

Notes: The error bars show the 95% confidence intervals using 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 6 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 7, we refine the analysis of Figure 6 by estimating Equation (41) 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 7 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 7 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
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)
log(resp. rate)	0.004	0.008	0.004	0.016^{**}	-0.001	0.011^{*}	0.001
	(0.004)	(0.006)	(0.004)	(0.006)	(0.007)	(0.005)	(0.005)
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, and the second-order polynomial of the number of ratings. Loc., Acc., Clean., and Comm. denote the location, accuracy, cleanliness, and communication rating, respectively.

Table 8: First-Differences Regression of Ratings on Low and High Prices

	Overall	Value	Loc.	Acc.	Clean.	Comm.	Check-in
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)
log(resp. rate)	0.004	0.008	0.004	0.016^{**}	-0.001	0.010^{*}	0.001
	(0.005)	(0.007)	(0.006)	(0.007)	(0.007)	(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. Note that the price coefficients show the slope for each price category. Standard errors are shown in parentheses and are clustered at the listing level. Each regression additionally controls for the month fixed effects and the second-order polynomial of the number of ratings. 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 (15). Additionally, we provide robustness checks for the results presented in Figure 2 when including further observable characteristics of the listings.

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

$$\begin{aligned} p_{it} =& \sum_{d=1}^{10} [\lambda_{d1} \mathcal{I}\{nr_{it} \in d\} \mathcal{I}\{r_{it} = 5\} + \lambda_{d0} \mathcal{I}\{nr_{it} \in d\} \mathcal{I}\{r_{it} < 5\}] \\ & + \mu_i + \mu_t + \epsilon_{it}, \end{aligned} \tag{42}$$

where p_{it} denotes the offer price of listing i in month t, nr_{it} and r_{it} are the number of ratings and the star rating at the *beginning* of month t, respectively, μ_i captures the listing fixed effects, and μ_t the month fixed effects, d denotes the deciles of the distribution of the number of ratings, λ_{d0} captures the conditional average price in a given decile of the number of ratings, d, for listings with *less* than five-star ratings, while λ_{d1} captures the conditional average price in a given decile of the number of ratings, d, for listings with a five-star rating.

After estimating Equation (42), we compute

$$q_{it} = \mu_i + \sum_{d=1}^{10} [\lambda_{d1} \mathcal{I}\{nr_{it} \in d\} \mathcal{I}\{r_{it} = 5\} + \lambda_{d0} \mathcal{I}\{nr_{it} \in d\} \mathcal{I}\{r_{it} < 5\}] \tag{43}$$

and obtain the quintiles of q_{it} based on which we create the five-scale indicator variable capturing the quality of a given listing, i, in period t.

Figure 8 repeats the analysis shown in Figure 2 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 9 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.[31] 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 (15) 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).

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

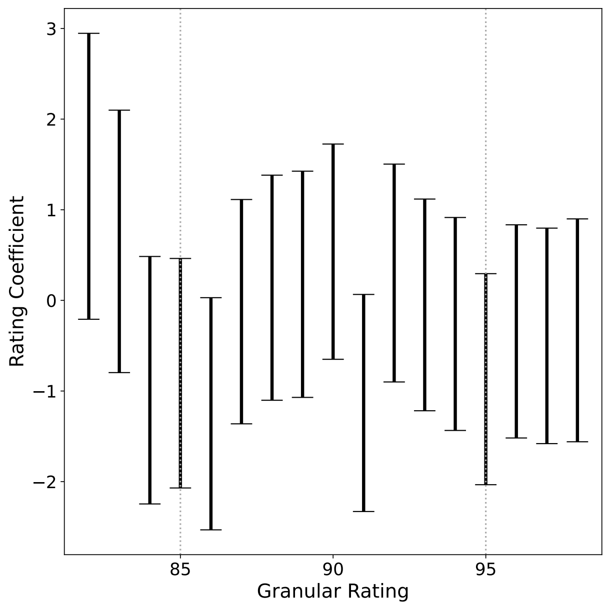


Figure 8: Conditional Average Offer Prices Around Salient Threshold

Notes: The error bars show the 95% confidence intervals using heteroskedasticity-robust standard errors.

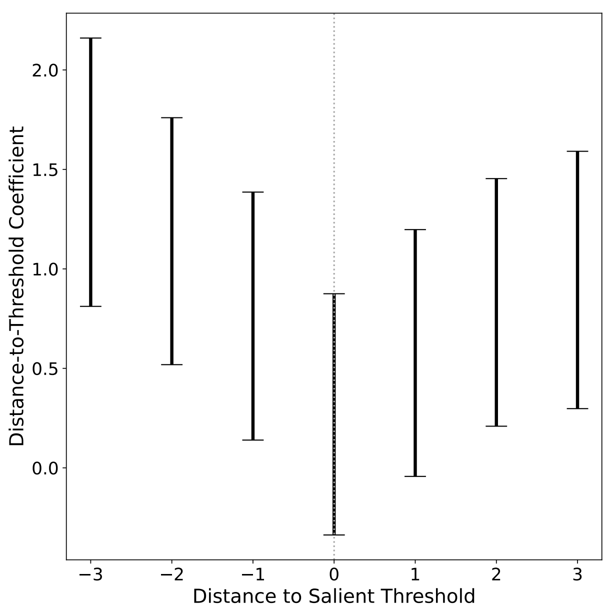


Figure 9: Conditional Average Offer Prices Around Salient Thresholds (When Pooling Thresholds)

Notes: The error bars show the 95% confidence intervals using heteroskedasticity-robust standard errors.

F Non-Parametric Estimation of Continuation Value

By estimating separate coefficients for each ratings level, Equation (44) provides a more flexible method to estimate the relationship between ratings and the monthly revenue compared to Equation (16). Figure 10 shows the OLS estimates of the coefficient β_j in Equation (44). The control variables and fixed effects used in Equation (44) are identical to those used in Equation (16) in the main text.

$$pq_{it} = \sum_{j=80}^{100} \beta_j \mathcal{I}\{r_{it} = j\} + X'_{it}\gamma + \mu_i + \mu_t + \epsilon_{it}\,. \tag{44}$$

Next, we estimate a local linear fit of the residuals obtained from estimating Equation (45). Figure 11 shows the results of this exercise. Again, the control variables used for the estimation are identical to those used in the main text. The local linear fit relies on a Gaussian kernel with a standard deviation of three.

Figure 10 and Figure 11 corroborate a convex-concave relationship between the monthly revenue and the granular rating measure.

$$pq_{it} = \beta_0 + X'_{it}\gamma + \mu_i + \mu_t + \epsilon_{it}\,. \tag{45}$$

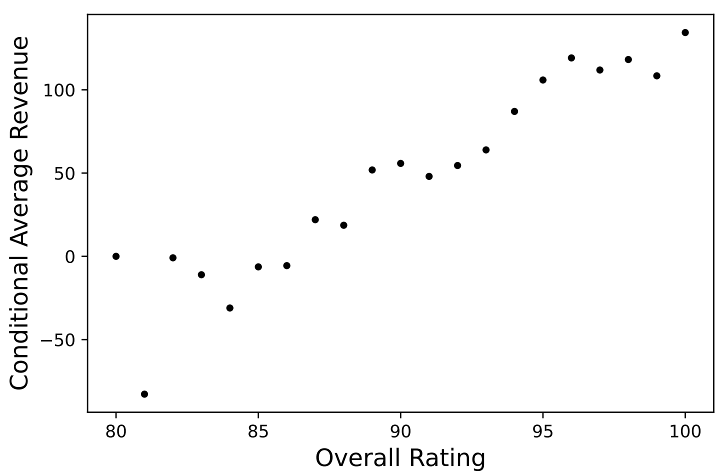


Figure 10: Relationship Between Monthly Revenues and Overall Ratings.

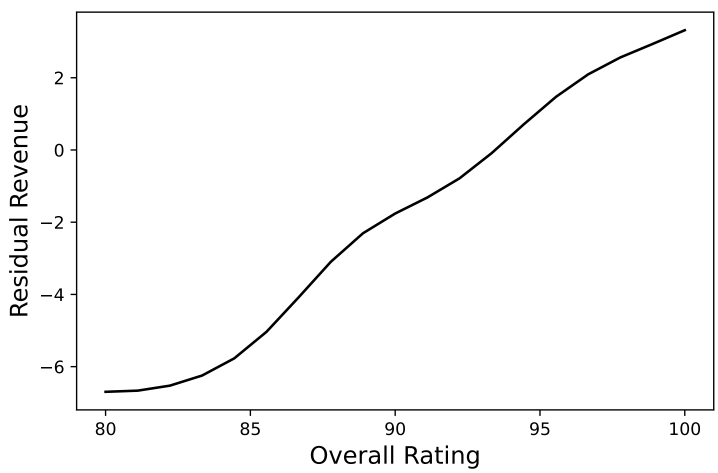


Figure 11: Local Polynomial Regression of Revenue Residuals on Overall Rating

G Price-Rating Dynamics

Figure 12a shows the per-period difference in the share of superhosts between hosts who offer an initial discount and hosts who do not. The superhost status is awarded by Airbnb as a function of specific performance metrics, like the average star rating, response rate, number of trip cancellations, etc.

While the metrics are regularly updated, it is likely that the eligibility criteria at the time the data were sampled were the following: (i) having hosted more than ten guests, (ii) maintaining a certain star average, (iii) staying below a certain trip-cancellation threshold, and (iv) maintaining a certain response rate.[32]

At the time of entry, the share of superhosts is similar across both groups but appears to diverge slowly over time. The superhost status can be interpreted as a measure of professionalization. However, it is important to note that, to the extent that the initial discount leads to higher ratings and drives demand, the higher share of superhosts could be a consequence of the initial discount.

Figure 12b shows the per-period difference in the share of hosts that activated the "instant bookable" option. This allows guests to book the accommodation without waiting for approval from hosts. In essence, this feature commits the host to accept any booking request and therefore offers a hotel-like booking experience. The share of instant booking features is higher within the group of hosts offering a discount. This provides supporting evidence that hosts offering a discount are more professional.

Figure 13 shows the conditional correlation between the initial discount (rescaled at its median) and the various rating subcategories as well as the logs of the bookings and revenues.

Including Listing Fixed Effects

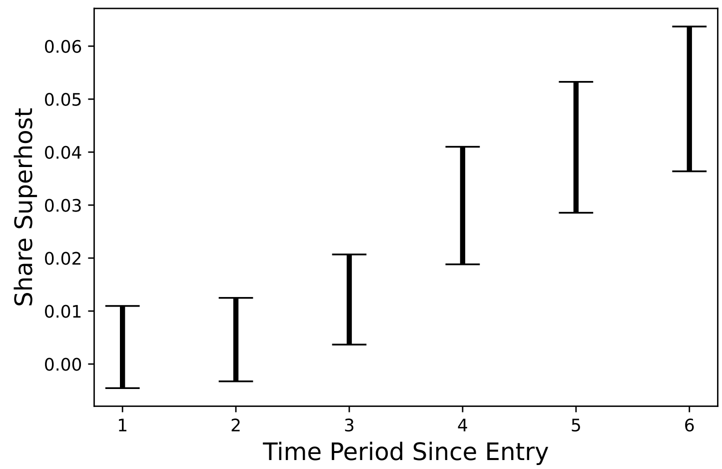
Because fp_i does not vary within listings, a fixed effect estimation of Equation (18) can only identify *changes* in the relationship between y_{it} and fp_i across periods – but not the relationship between y_{it} and fp_i in the baseline period (which we define as the first period).

Figure 14 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 effects of fp_i on the changes of the dependent variables are very consistent with the results of the main analysis.

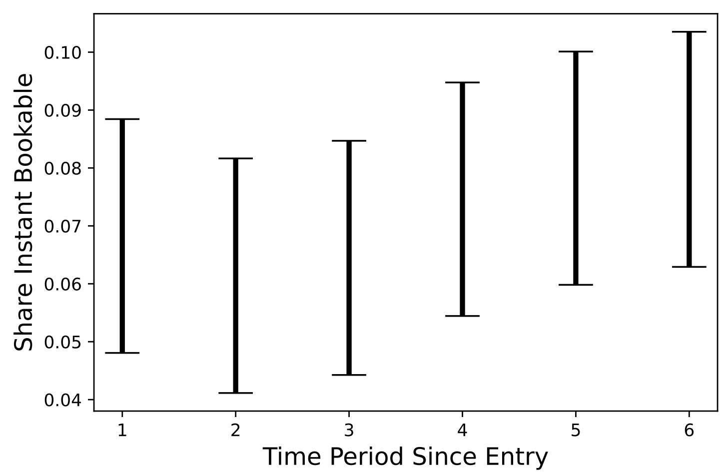
In other words, the changes shown in Figure 14 correspond well to the changes

[32] See https://airhostsforum.com/t/superhost-when-do-i-really-meet-all-the-requirements/5603/4 (last accessed: August 16, 2022).

observed in Figure 4 between the coefficients for $\tau \geq 2$ and $\tau = 1$. This mitigates concerns related to endogeneity stemming from unobserved time-constant listing covariates.



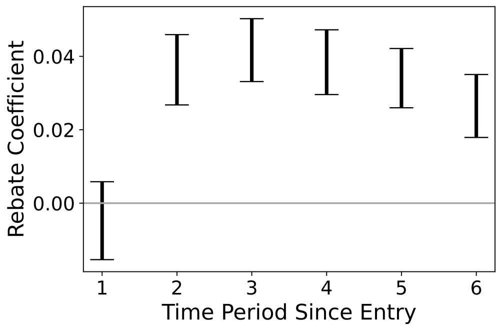
(a) Share Superhosts



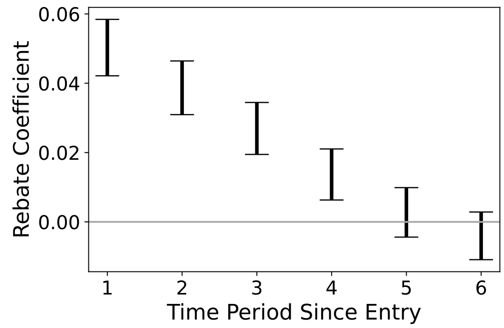
(b) Share Instant Bookable

Figure 12: Relationship Between Measures of Host Professionalization and Discount

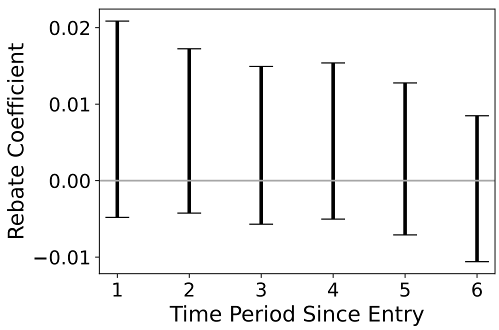
Notes: The error bars show the 95% confidence intervals using standard errors clustered at the listing level.



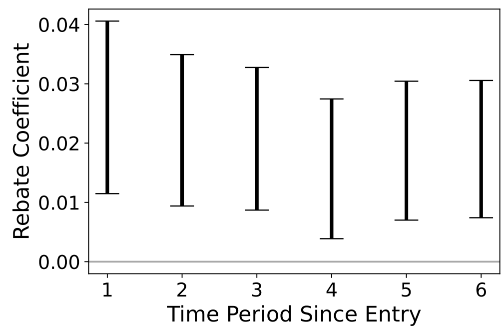
(a) y_{it} = Revenue (log)



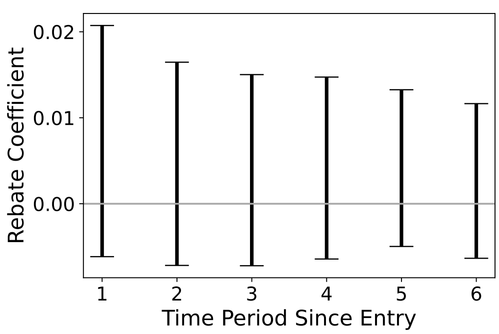
(b) y_{it} = Number of Bookings (log)



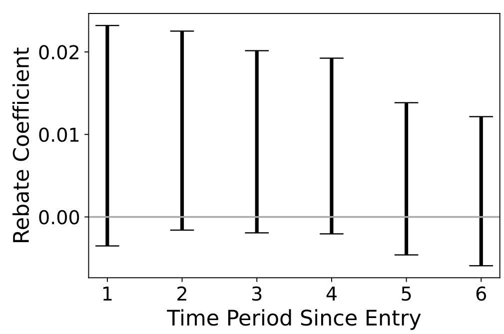
(c) y_{it} = Accuracy rating



(d) y_{it} = Cleanliness rating



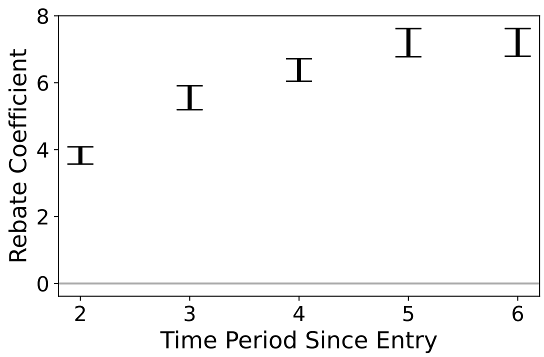
(e) y_{it} = Communication rating



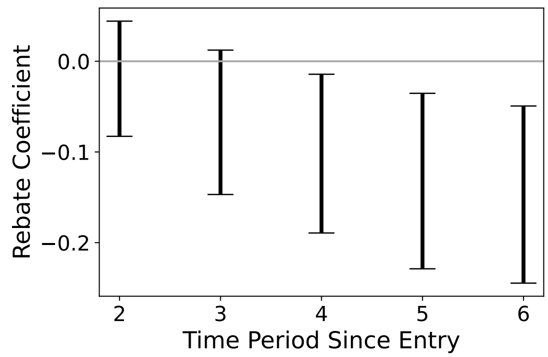
(f) y_{it} = Check-in rating

Figure 13: β_τ for Different Dependent Variables (Scaled at Median Discount of 7%).

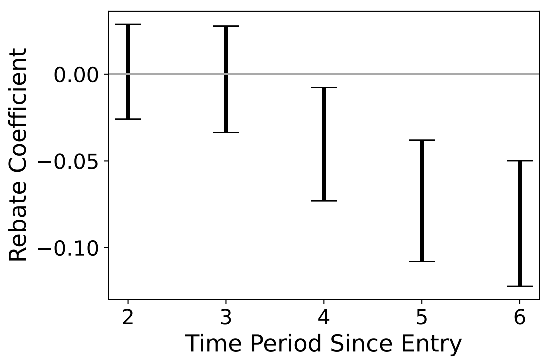
Notes: The error bars show the 95% confidence intervals using standard errors clustered at the listing level.



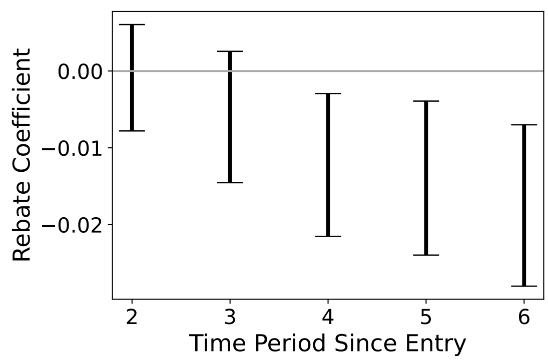
(a) y_{it} = Booking Price



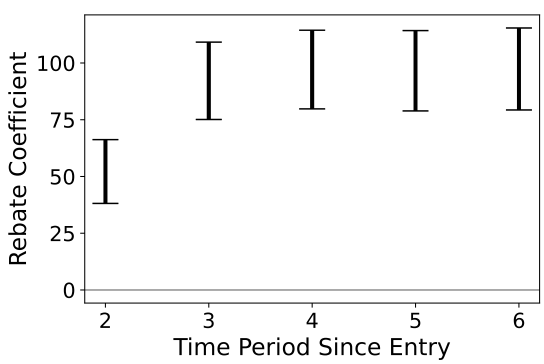
(b) y_{it} = Overall Rating



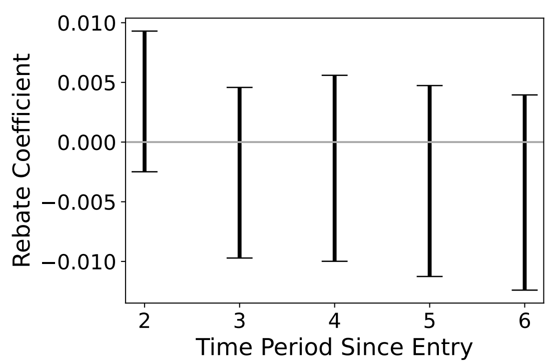
(c) y_{it} = Number of Bookings



(d) y_{it} = Value-for-Money rating



(e) y_{it} = Revenue



(f) y_{it} = Location Rating

Figure 14: β_τ for Different Dependent Variables (Scaled at Median Discount of 7%)

Notes: The results are obtained form regressions including listing fixed effects. The error bars show the 95% confidence intervals using standard errors clustered at the listing level.