Value for Money and Selection: How Pricing Affects Airbnb Ratings

Christoph Carnehl
Bocconi University
christoph.carnehl@gmail.com

Maximilian Schaefer
University of Bologna
maximilian.schaefer@unibo.it

André Stenzel
University of Mannheim
andre.stenzel@uni-mannheim.de

Kevin Ducbao Tran
University of Bristol
kevin.tran@bristol.ac.uk

July 22, 2021

Abstract

We investigate the impact of prices on seller ratings. In a stylized model, we illustrate two opposing channels through which pricing affects overall ratings and rating subcategories. First, higher prices reduce the perceived value for money which worsens ratings. Second, higher prices increase the taste-based valuation of the average traveler which improves ratings. Using data from Airbnb, we document a negative relationship between prices and ratings for most rating subcategories indicating that the value-for-money effect dominates the selection effect. In line with our model, we find that hosts of low-rating listings exert more effort than those of high-rating listings. Finally, an empirical assessment of the dynamics in the market suggests that taking the effect of prices on future ratings into account pays off: entrants who set low entry prices obtain better ratings and higher revenues in the medium run. A median entry discount of 8.5 percentage points increases medium-run monthly revenues by approximately 50 euros.

1 Introduction

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

In this paper, we use an illustrative theoretical model to derive predictions regarding strategic pricing incentives of sellers. Subsequently, we assess our model predictions empirically using a unique transactions and ratings dataset of Airbnb listings in Paris, France, in 2017. Our data allow us to determine when listings were booked and when they received a review. We use this information to match the price of a booking to reviews. First, we show that higher prices are associated with lower ratings. Consistent with our model predictions, we next provide evidence that hosts charging relatively lower prices when entering the platform benefit from their early discounts in the medium run in terms of both ratings and revenues. A discount of 8.5 percentage points in the first period increases medium run monthly revenues by approximately 50 euros compared to listings which do not offer discount when entering the market.[1]

In our model, a long-lived host offers an apartment of fixed quality to short-lived travelers.[2] In each period, the host chooses the price of the listing as well as an effort level. Prospective travelers observe two types of aggregate ratings: a value-for-money rating and an effort rating. Based on these ratings, they build expectations about the quality of the listing and the host's effort. Travelers' decisions to book depend on these expectations, their idiosyncratic preference for the listing, and the price.

Travelers who stay at the apartment provide ratings that depend on the apartment's quality, the host's effort, their own idiosyncratic taste, and the

[1]The average monthly revenues in our sample amount to approximately 587 euros.

[2]The model is similar to Stenzel et al. (2020), but explicitly incorporates hosts' effort decisions.

price. We assume that higher quality, effort, and taste induce a higher value-for-money rating. Further, we assume that a higher price lowers the value-for-money rating directly, but can increase the rating because the average traveler who decides to book the apartment will have a higher preference for it. 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.[3] The net effect depends on the relative importance of the value-for-money and the selection effect. For the effort rating, we assume that higher prices always have a negative impact.

Importantly, the hosts' current pricing and effort decision have an impact on the ratings left by current-period travelers and thereby future profits. We use the model to derive testable hypotheses on the relationship between prices and rating scores, as well as the prices and effort.

Our first set of empirical results indicates that, in the context of our data, the value-for-money effect is stronger than the selection effect. We find that higher prices are generally associated with worse ratings. This relationship is particularly pronounced for the value-for-money rating and disappears for the location rating. These results are consistent with the mechanisms at play in our model. The negative relationship between prices and the value-for-money rating suggests that the value-for-money effect dominates the selection effect in general. The absence of this relationship for the location rating in turn 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 for a very specific taste component of the traveller.

Decomposing the analysis with respect to above- and below-median priced listings, we find that the negative relationship is stronger for low-price listings. This result is in line with a higher price sensitivity of travelers choosing cheaper accommodations which, in the context of our model, increases the relative importance of the value-for-money effect. In line with the baseline results, this finding is strongest for the value-for-money rating and absent for the location rating.

Another prediction of our model pertains to the strategic relationship between a host's price and effort decision, which can be either strategic complements or substitutes. Whether there is complementarity or substitutabil-

[3]The reason for the selection effect is straightforward: all else equal, a marginal traveler who was indifferent before the price increase will no longer be willing to book the apartment. As the marginal traveler is the traveler with the lowest preference for the apartment, the price increase thus leads to a higher average preference for the apartment.

ity between prices and effort depends on the shape of the function of future profits in the ratings, which may vary with the level of ratings. In particular, if a rating increase leads to a higher marginal benefit of a further increase in the rating (i.e. if the future profits are convex in the rating), price and effort are strategic complements.

We first investigate the relationship between revenues and ratings by regressing monthly listing revenues on a third-degree polynomial of the overall star-rating as well as other covariates. This specification captures a flexible relationship between the rating and revenues, which are in turn an essential component of the continuation value. The results suggest a convex relationship for lower ratings and a concave relationship for higher ratings.[4] Given this shape of the revenue function, the model predicts effort and prices to be complements while a listing has lower ratings and substitutes for higher ratings. Our empirical exercise confirms this prediction: when we regress effort ratings on prices, we find a negative relationship for listings with low overall rating and no relationship for listings with high overall ratings.

Finally, the model suggests that hosts can use strategic pricing when entering the platform to obtain better ratings which they can benefit from in future periods. The underlying reason is that prices can be adjusted frequently but have a persistent effect on the ratings which determine future demand. In particular, hosts should enter with a price discount relative to a naive entry price ignoring this dynamic effect of entry prices if the value-for-money effect dominates the selection effect.

Again, this theoretical hypothesis is supported empirically. We find that new hosts who charge a comparatively lower price when entering the platform receive better value-for-money ratings and more bookings in the early periods. This allows these hosts to charge relatively higher prices in subsequent periods resulting in higher revenues in the medium run. A discount of 8.5 percentage points in the first period increases medium run monthly revenues by approximately 50 euros compared to listings which do not offer a discount when entering the market

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

[4]This shape seems intuitive also from the perspective of (Bayesian) belief-adjustments. Upward adjustments to beliefs due to a positive signal are likely to be stronger for intermediate beliefs compared to when the traveler was almost certain that the apartment is of high value.

quantities (Livingston, 2005).

In contrast, we are interested in the opposite mechanism: whether prices affect ratings. As such, we add to the literature on strategic rating management through price setting. While our main contribution is empirical, our theoretical model is closely related to Stenzel et al. (2020), 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 serves as an illustration of 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. Zegners (2019) finds that books offered for free on an online self-publishing platform generate more, but worse reviews. The author argues that this result is due to a selection effect in which readers who read a free book have a lower preference for it. This insight is inline with the selection effect in our model. We add to the paper by considering continuous variation in prices rather than a comparison between a zero price and positive prices.[5] Furthermore, we consider an additional effect that prices can have on ratings: the value-for-money effect. Indeed, we find that this value-for-money effect 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 result in a decrease in average rating. Our theoretical model provides an explanation for this result: a dominant value-for-money effect. Sorokin (2021) finds that producers on the video game platform Steam use discounts to transition to higher review tiers. The author suggests two potential mechanisms: First, consumers who buy during a discount leave better reviews. This result is in line with the value-for-money effect we have in mind. Second, when a game is close to being upgraded in their review tier, better reviews move the game to the upper tier, whereas worse reviews have no downside, as the game just remains in its review tier. Therefore, discounts are a useful tool to increase the number of reviews, even given that these reviews will arrive with some variance in their value. Jointly, our paper contributes to this literature by proposing a unified theoretical framework that can explain all of these empirical results. Our empirical part provides evidence in line with this framework.

Our paper is also related to research on the determinants of ratings more broadly. Cabral and Li (2015) find that lower quality transaction result in

[5] 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).

more negative feedback. Mayzlin et al. (2014) find evidence of hotels faking negative reviews for their competitors and positive reviews for themselves on TripAdvisor. Luca and Zervas (2016) provide evidence of restaurants using fake reviews on Yelp and He et al. (2021) study the market for fake product reviews 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 reviews, hotels tend to receive fewer, but longer negative reviews. We add to this literature by highlighting both theoretically as well as 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 datasets and provide some descriptive statistics. In Section 4, we describe our main empirical analysis and results. In Section 5, we turn our attention to an analysis of strategic price-setting by hosts entering the platform. 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. We outline each component of the model in more detail below.

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 parametrizes the initial attitude of potential travelers towards the apartment, i.e., 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 and effort $e \geq 0$ at a cost $c(e) = c/2e^2$. We assume that the host maximizes long-run profits and has a discount factor $\delta \in [0, 1]$.

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.[6]

[6]It is straightforward to also include an overall rating, Ψ^o, into the model setup. For the purpose of our analysis, we abstract from it 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

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 quality of the apartment, the effort of the host in the period of their stay, their idiosyncratic taste for the apartment, and money. We assume that 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. However, they have access to the rating Ψ.

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 given by $\mu(\Psi^v) \in [0,1]$ with $\mu(0) = 0$ and $\mu'(\Psi^v) > 0$; i.e., 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 of the apartment. Similarly, we assume that the belief about the host's exerted effort level is $\nu(\Psi^e) \in [0,1]$ with $\nu(0) = 0$ and $\nu'(\Psi^e) > 0$. Thus, better effort ratings induce higher effort beliefs.

We assume that travelers are risk neutral and have an additively separable utility function

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

and an outside option which we normalize to zero.

It follows that a traveler books the apartment if $u(\Psi^v, \Psi^e, \omega_i, p) \geq 0$, which implies that there is an indifferent 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 taste parameters are uniformly distributed, we can immediately derive the demand function[7]

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

Note that in our setup the host's effort choice does not affect flow profits. However, it nonetheless matters for future profits as it affects the updated ratings available to future travelers, which we explain in detail next.

be qualitatively similar if travelers alternatively only observe an aggregate rating Ψ^o which is a (known) function $f(\Psi^v, \Psi^e)$.

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

Rating generation Ratings are generated by visitors of the apartment, that is, by travelers who book and stay at the apartment. We assume that travelers rate non-strategically and in particular that every traveler that visited an apartment rates it with a fixed probability that is independent of any traveler-apartment-specific characteristics.[8] 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 effort by the host exerted in the period of the traveler's stay e, the traveler's idiosyncratic taste ω_i, and the price paid for the stay p,[9]

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

Thus, a higher quality, effort, and taste induce a higher value-for-money rating while a higher price induces a lower value-for-money rating.[10] 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 staying at the apartment on top of this direct effect.

We assume that 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.

We assume—without loss of economic insight—that only the average rating left by travelers staying at the apartment is used to update the ratings displayed to future travelers. This directly implies that this average rating is deterministic, which simplifies the exposition. Note that 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. In the present setup, the average *purchasing* traveler's taste is used to compute the

[8] 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 those discussed in Stenzel et al. (2020).

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

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

average value rating left in any given period, and is given by

$$\begin{aligned}\omega^e &= \frac{1+\tilde{\omega}(\Psi^v, \Psi^e, p)}{2} \\ &= \frac{1+p-\mu(\Psi^v)-\nu(\Psi^e)}{2}. \end{aligned} \tag{5}$$

This allows us to write the induced average value-for-money rating as

$$\psi^v(\theta, e, \omega^e, p) = \varphi^v\left(\theta, e, \omega^e(p), p\right), \tag{6}$$

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

Ratings are updated from one period to the next as a simple average.[11] 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

$$\Psi_{t+1}^j = \frac{t}{t+1}\Psi_t^j + \frac{1}{t+1}\psi_t^j. \tag{7}$$

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 attitude towards the listing based on the description of the apartment on the platform.

Myopic host As a benchmark, we consider a myopic host who simply maximizes flow profits. Such a host maximizes

$$(1+\mu(\Psi_t^v)+\nu(\Psi_t^e)-p)p - \frac{c}{2}e^2 \tag{8}$$

with respect to p and e. This immediately gives

$$p_t^m = \frac{1+\mu(\Psi_t^v)+\nu(\Psi_t^e)}{2}, \quad e_t^m = 0, \tag{9}$$

i.e., that a myopic host never exerts effort and chooses the monopoly price given current beliefs of travelers.

[11] Airbnb does not explicitly state whether individual ratings are aggregated using simple averages, but our analysis suggests that this is the case, see Figure 1.

The host's problem The strategic host takes the effect of current actions on future profits into account. Thus, the host maximizes the discounted sum of profits

$$\max_{(p_t,e_t)_{t=1}^T} \sum_{t=1}^{T} \delta^t \left(q(\Psi_t^v, \Psi_t^e, p_t)p_t - \frac{c}{2}e_t^2 \right) \tag{10}$$

subject to the law of motion of the ratings defined above. It is convenient to invoke the dynamic maximum principle and to rewrite the optimization problem as a Bellman equation[12]

$$V_t(\Psi_t^v, \Psi_t^e) = \max_{p_t,e_t} \left(q(\Psi_t^v, \Psi_t^e, p_t)p_t - \frac{c}{2}e_t^2 + \delta V_{t+1}\left(\Psi_{t+1}^v(p_t,e_t), \Psi_{t+1}^e(p_t,e_t)\right)\right). \tag{11}$$

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

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

The impact of the price on the value-for-money rating depends on the relative strength of the value-for-money effect and the selection effect:

$$\frac{d}{dp_t}\psi^v > 0 \iff \frac{-\varphi_p^v}{\varphi_\omega^v} < \frac{1}{2}. \tag{12}$$

We relegate all derivations to Appendix A. The impact of effort on both ratings, and of the price on the effort rating are by assumption on the rating function. The condition (12) in turn obtains by noting that a change in price affects the value-for-money rating both via the direct impact on the induced rating (φ_p^v), and via changing the taste of the average traveler ($\varphi_\omega^v \cdot \frac{d\omega^e}{dp} = \frac{1}{2}\varphi_\omega^v$). The relative size of the direct and indirect impact, which go in opposite directions, thus determines 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

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

rating is increasing in the price, while 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 her rating. It is immediate that she has an incentive to exert effort as to positively influence the future rating stock. The effect on the pricing incentives, however, is again ambiguous and depends in particular on the relative strength of the value-for-money and selection effect, but also on the extent to which the price impacts the effort rating.

Lemma 2 *A strategic host exerts effort in every period. Whether a strategic host prices above or below the myopically optimal price depends on the sign of*

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

Equation (13) reflects that the price impacts not only 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 impact the price has on the effort rating. Note that whenever $\varphi_p^e = 0$, i.e., when the price *only* affects the value-for-money rating, the sign of (13) 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 price and selection effect. If this is not the case, there is the additional detrimental impact of a price increase on the effort rating, which all else equal strengthens incentives to lower prices for rating management.

Prices and effort—substitutes or complements? A naturally arising question is whether and how hosts trade off prices and effort in their rating management. We can address this by analyzing how the optimal price and effort levels co-vary. We say effort and price are strategic complements if an increase in effort induces a further distortion of the price from the myopic price; i.e., an increase in effort increases the incentive to use the price to improve the rating even further.

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

$$sign\left(\frac{d^2V_{t+1}}{de_t dp_t}\right) = sign\left(\frac{dV_{t+1}}{d\Psi^v}\left(\varphi_\omega^v \frac{1}{2} + \varphi_p^v\right) + \frac{dV_{t+1}}{d\Psi^e}\varphi_p^e\right). \tag{14}$$

In this case, effort provision increases the incentives to distort the price away from the myopically optimal price. Otherwise, price and effort are substitutes.

If $\varphi^e = 0$ and $\varphi^v_{pe} = 0$, strategic pricing and effort are complements if and only if

$$\frac{d^2 V_{t+1}}{d(\Psi^v)^2} + 2\frac{d^2 V_{t+1}}{d\Psi^v d\Psi^e} > 0. \tag{15}$$

The economic forces driving the complementarity/substitutability of strategic pricing and effort provision are best exemplified by muting the price impact on the effort rating, and cross-effects in the value-for-money rating. In this case,

$$\frac{dp_t}{de_t} = \frac{\delta}{2(t+1)^2}\frac{d\psi^e}{de_t}\frac{d\psi^v}{dp_t}\left(\frac{d^2 V_{t+1}}{d(\Psi^v)^2} + 2\frac{d^2 V_{t+1}}{d\Psi^v d\Psi^e}\right), \tag{16}$$

so that $\frac{dp_t}{de_t}$ has the same sign as $\frac{d\psi^v}{dp_t}$ provided that (15) is satisfied. Note that this is the case whenever the continuation value V_{t+1} is sufficiently convex.

One way in which (15) can become negative, rendering strategic pricing and effort provision substitutes instead of complements, is when the continuation value is concave in the value-for-money rating. Intuitively, this can arise at high beliefs, so that further increases in the belief that the good is of high quality have little additional impact on prices or demand. More specifically, if the curvature of the continuation profits exhibit a convex-concave shape, i.e., are convex for low rating levels, but concave for high rating values, hosts have an incentive to drive up their ratings early on using both instruments, price and effort, in conjunction as strategic pricing and effort are complements. However, once the ratings and associated beliefs surpass a certain threshold, they may slack on one of the two instruments and make up for it with the other one because they have become substitutes.

Importantly, whether strategic pricing and effort are substitutes or complements does not itself 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 selection effect is dominant. For example, strategic pricing and effort being complements would imply a negative relationship between prices and efforts—lower prices are associated with increased effort levels—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, and the latter leads to downward pricing pressure because the dominant value-for-money effect implies that lowering prices increases the rating obtained in the future.

Hypotheses for the Empirical Analysis The theoretical analysis gives rise to several hypotheses which test either 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 at play. Specifically, we obtain hypotheses about the role of effort, the role of 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 effort.*

Hypothesis 1 essentially comprises the model assumptions that host effort positively impacts both the value-for-money and the effort rating.

With respect to the impact of the price on the respective ratings, Lemma 1 offers potentially competing hypothesis. 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 effect. 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 on Yelp.

Hypothesis 2 (Role of Price) *Both the value-for-money rating and the effort rating are negatively affected by 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 behavior—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 behavior 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 value" at particular rating levels. We consider it plausible that the marginal benefit of rating increases is decreasing for sufficiently high ratings, but increasing otherwise. This would imply that strategic pricing and effort provision are complements for low rating stocks, but become substitutes once the host has induced 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 prices 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 sophisticated hosts overall charge lower prices than unsophisticated sellers because the value-for-money effect dominates.

Hypothesis 4 *Sophisticated hosts charge lower prices than unsophisticated hosts conditional on the current rating stock. However, sophisticated hosts charge higher prices and maintain a higher rating than unsophisticated hosts after sufficiently many traveler stays which render the ratings less responsive.*

Note that within our model, this hypothesis would directly follow from Hypothesis 2—dynamically, the host further internalizes the weakly negative impact of the price on the effort rating, which provides a further incentive for lower prices from a dynamic perspective. The second part of Hypothesis 4 is intuitive once observing the decreased impact of incoming ratings by recent travelers on the aggregate rating. Sophisticated hosts charge lower prices within-period than unsophisticated hosts to positively impact the rating stock. Such strategic behavior implies that their rating will reach a higher value than that of unsophisticated hosts over time. As the rating becomes sufficiently unresponsive over time, the dynamic pricing incentives lessen, 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 of 2017.

The transactions data were provided by Airdna, a specialist for vacation rental data. The Airdna data allow us to determine for each listing whether it was available or booked on a particular date, as well as the corresponding asked daily price. Consecutive days of occupancy by the same guests are identified by booking identifiers.

Table 1: Summary Statistics

	count	mean	min	p50	max
Days Available	493488	18.61	0.00	24.00	39.00
Days Booked	493488	10.83	0.00	7.00	39.00
# Total Bookings	493569	2.74	0.00	1.00	31.00
Price (All Bookings)	316366	100.24	9.00	78.00	6000.00
# Reviewed Bookings	383510	1.24	0.00	0.00	22.00
Price (Reviewed Bookings)	183893	95.94	9.00	75.64	2500.00
Overall Star Rating	493569	4.70	1.00	5.00	5.00

Notes: All variables have been aggregated on the "monthly" level. For example, the number of available days is calculated as the average number of days a listing was available for booking between two consecutive rating updates. The time frame between two consecutive rating updates can exceed one month, depending on the scraping routine. This variation explains why the maximum number of available days is 39 and not 31.

The ratings data were provided by InsideAirbnb.com. The data contain monthly updates of the aggregate star ratings in various rating categories. The updates are observed at the beginning of each month. We observe star ratings in six 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 the overall experience of a guest during a stay in a particular listing. It is the rating which is displayed most prominently to potential guests on the Airbnb website. The other rating subcategories capture specific aspects of the stay. These subcategories are only seen by guests that browse the accommodation more thoroughly. It should be noted that the overall rating is not a mechanical average of the rating subcategories but can be freely chosen by the customer.

Because we observe the ratings only in monthly intervals, we cannot directly match transaction prices to corresponding ratings. Instead, we analyze the impact of the average prices charged in a month on the evolution of the aggregate ratings. To discard bookings which did not receive a rating, we make use of review timestamps which allow us to determine the date on which a review was submitted for a particular listing (but not the associated rating). Guests can leave a review on Airbnb within 14 days of 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 review, we choose the closest one. For each listing, we calculate the average price of the rated bookings in each month.

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 18 days and booked for ten days per month.[13] During this period, there are on average 2.74 bookings of which 1.24 are reviewed.[14] The average price of reviewed bookings is 95.94 euros per night. The summary statistics for the overall star rating reveal that a majority of the observations enjoy the highest possible rating.

Airbnb itself does not disclose exactly how individual reviews enter into the displayed aggregate rating. However, the data suggest that the aggregate rating is obtained by simple averaging of individual ratings, which is in line with anecdotal evidence.[15] Figure 1 shows the probability of observing a change in the aggregate overall rating between two consecutive monthly

[13]Note that the time between two rating updates is not always exactly one month, but can vary due to the scraping procedure.

[14]The percentage of reviewed bookings is reported to be equal to 68 percent in the study of Fradkin et al. (2020). However, the geographical scope of their data is not specified. Estimated review rates across large cities are substantially lower, see, for example, http://insideairbnb.com/get-the-data.html (last accessed: July 19, 2021).

[15]See, for example, hhttps://airhostsforum.com/t/how-exactly-is-the-star-rating-calculated/14575m (last accessed: July 02, 2021) for an interesting discussion among Airbnb hosts on how Airbnb ratings are aggregated over time. The discussion in https://community.withairbnb.com/t5/Hosting/5-stars-in-all-categories-but-4-star-stay/td-p/6934705 (last accessed: July 02, 2021) clarifies that the overall rating and the ratings in the subcategories are, in principle, independent. Therefore, the overall rating does not appear to be a weighted average of the subcategory ratings.

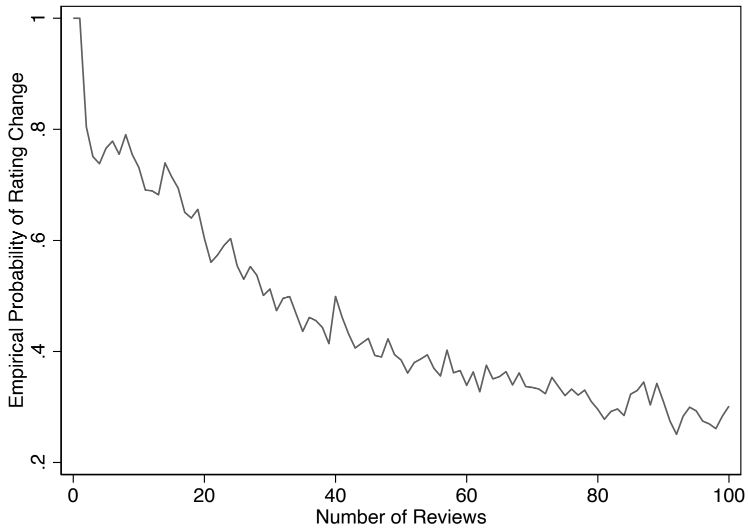


Figure 1: Empirical Probability of Observing a Rating Change in the Overall Rating.

updates as a function of the number of reviews a listing 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 which received more ratings in the past.

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 begin by providing evidence that the value-for-money effect dominates the selection effect for most rating subcategories and the overall rating.[16]

[16]Our theoretical and empirical analysis builds on the assumption that ratings affect demand; otherwise, our insights are trivial. Establishing this link between ratings and demand is not our main focus and, therefore, we do not present results pertaining to the

4.1 Value-for-Money vs. Selection Effect

Figure 2 shows the cumulative distribution of the overall star rating for listings in different price terciles. The price terciles of listings were computed based on the average prices observed for each listing. Figure 2 reveals that listings in a lower price tercile have systematically lower ratings as compared to listings in higher price terciles. For example, 45 percent of the listings in the highest price tercile have the highest possible rating. In the lowest price tercile, only 30 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 price. 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.

It should be clear that a more detailed analysis which addresses these concerns is necessary. In our empirical analysis, we address the first potential explanation by controlling for unobserved time-constant quality differences between listings by including listing-specific fixed effects. To address reverse causality, we use our matched rating-price pairs and regress period t-ratings on period $t-1$-prices. We argue that after taking into account these two potential explanations, the remaining conditional correlation between ratings and prices can be interpreted as the relative importance of the value-for-money and the selection effect.

Specifically, we estimate the following equation

$$rat_{it}^{cat} = \beta_0 + \beta_1 \times log(p_{it-1}) + X_{it}'\gamma + \mu_i + \mu_t + \epsilon_{it}. \tag{17}$$

rat_{it}^{cat} denotes the aggregate star rating for category cat of listing i at the start of month t. $log(p_{it-1})$ denotes the 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 rating we observe immediately after the transactions took place. This time lag between prices and ratings helps address reverse causality concerns. The listing fixed

impact of ratings on demand in the main text. 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 B, we present descriptive and causal evidence using our data that is consistent with this finding.

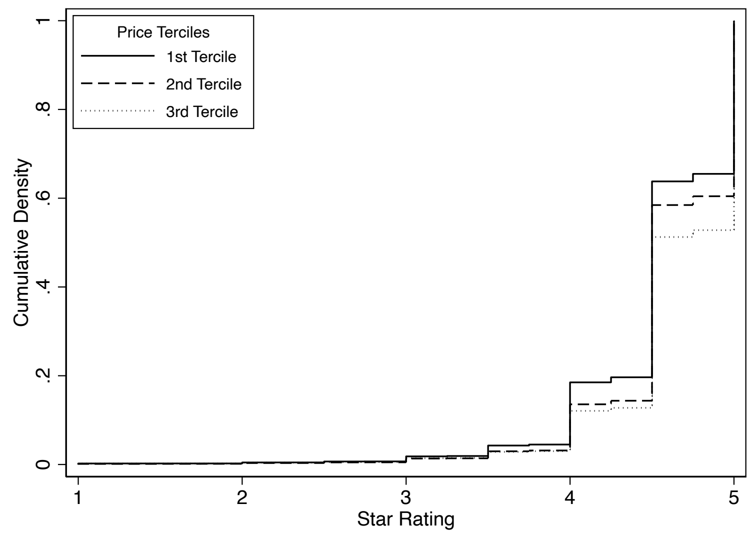


Figure 2: Rating Distribution for different price terciles

.

effects μ_i account for time-invariant quality differences, while X_{it} accounts for time-variant factors, such as observed measures of host effort.

To control for host effort, we include information on the host response rate (i.e., how often does the host respond to inquiries of potential guests) and the host response time (i.e., how quickly does the host respond to guest inquiries). The first variable is a share between zero and one. The second variable is an indicator that takes on the value one if the average host response time is less than one hour. To account for the averaging in the calculation of the aggregate ratings, we also include the number of reviews and its square in X_{it}. Finally, μ_t denotes month fixed effects.

Table 2 shows the results obtained when estimating Equation (17) for each of the seven rating categories. Note that all regressions control for the number of reviews, month fixed effects, and listing fixed effects. The price coefficient is negative across all rating categories.

The results are intuitive in light of our model. The price coefficient is largest for the value-for-money rating, which is unsurprising. All other rating categories appear less affected by prices. Interestingly, the location rating, which asks for an arguably time-invariant quality aspect of the flat, is the least correlated with prices. Higher effort impacts ratings (weakly)

Table 2: Regression Results

	(1) Overall	(2) Value-for-money	(3) Location	(4) Accuracy	(5) Cleanliness	(6) Communication	(7) Check-in
Price (log)	-0.33***	-0.50***	-0.08	-0.33***	-0.02	-0.19**	-0.14*
	[-0.48,-0.18]	[-0.66,-0.35]	[-0.21,0.04]	[-0.47,-0.19]	[-0.18,0.14]	[-0.31,-0.06]	[-0.27,-0.01]
Host response rate (log)	0.31**	0.19	0.10	0.23*	0.23*	0.23**	0.12
	[0.11,0.50]	[-0.01,0.39]	[-0.06,0.26]	[0.05,0.42]	[0.02,0.43]	[0.06,0.39]	[-0.04,0.29]
Avg. response time $< 1h$	-0.02	-0.01	0.01	0.03	0.01	0.04	0.03
	[-0.07,0.03]	[-0.06,0.05]	[-0.03,0.05]	[-0.02,0.08]	[-0.04,0.07]	[-0.00,0.08]	[-0.01,0.08]
Observations	181730	181477	181493	181652	181718	181615	181517

Notes: *, **, *** indicate statistical significance at the five, one, and 0.1 percent level, respectively. The square brackets show 95 percent confidence intervals.

positively. Crucially, the overall rating, which is the most salient rating to customers, is negatively affected by the prices. This result suggests that the value-for-money effect indeed dominates the selection effect in our data as hypothesized.

Table 3: Regression Results - Price Interaction

	(1) Overall	(2) Value-for-money	(3) Location	(4) Accuracy	(5) Cleanliness	(6) Communication	(7) Check-in
Price (log) - Below mean	-0.43***	-0.65***	-0.12	-0.41***	-0.13	-0.12	-0.15
	[-0.67,-0.19]	[-0.90,-0.40]	[-0.32,0.08]	[-0.64,-0.18]	[-0.39,0.12]	[-0.33,0.08]	[-0.36,0.06]
Price (log) - Above mean	-0.27**	-0.42***	-0.06	-0.29**	0.05	-0.22**	-0.13
	[-0.45,-0.08]	[-0.61,-0.22]	[-0.22,0.09]	[-0.46,-0.11]	[-0.15,0.25]	[-0.38,-0.07]	[-0.29,0.03]
Host response rate (log)	0.31**	0.19	0.10	0.23*	0.23*	0.23**	0.12
	[0.11,0.50]	[-0.01,0.39]	[-0.06,0.26]	[0.05,0.41]	[0.02,0.43]	[0.06,0.39]	[-0.04,0.29]
Avg. response time $< 1h$	-0.02	-0.01	0.01	0.03	0.01	0.04	0.03
	[-0.07,0.03]	[-0.06,0.05]	[-0.03,0.05]	[-0.02,0.08]	[-0.04,0.07]	[-0.00,0.08]	[-0.01,0.08]
Observations	181730	181477	181493	181652	181718	181615	181517

Notes: *, **, *** indicate statistical significance at the five, one, and 0.1 percent level, respectively. The square brackets show 95 percent confidence intervals.

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 the average detrended prices of listings. The results in Table 3 are consistent with a higher price-sensitivity of economy customers, which, according to our model, should reinforce the value-for-money effect of prices on ratings.

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 do suggest that the value-for-money effect dominates the selection effect after accounting for time-invariant quality features of Airbnb listings. Based on our theoretical framework, this insight allows us to test the hypotheses on the price-effort relationship, as well as dynamic price and rating patterns, as it determines the predicted sign of the relationship 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, i.e., $\frac{d\psi^v}{dp_t} < 0$ in terms of our model. As a consequence, sellers are able to obtain higher ratings and hence future profits at the expense of reduced flow profits by strategically lowering the price of the listing. This insight is important when testing Hypothesis 3. The hypothesis posits that strategic pricing and effort are complements for low levels of the value-for-money rating, and substitutes for very high levels. This hypothesis builds on the assumption that the continuation value is convex in the ratings for low rating levels and becomes concave for high levels.

To test the hypothesis, we proceed in two steps. We first show that the continuation profit as a function of the value-for-money rating displays a convex-concave relationship. In a second step, we explicitly analyze the relationship between equilibrium prices and efforts. In light of the dominant value-for-money effect and the convex-concave continuation profits, the model predicts a negative relationship between prices and ratings for low levels of the value-for-money rating—with higher effort, the convexity implies an increased marginal benefit from strategic pricing to increase future ratings, which leads to downward pricing pressure in light of the dominant value-for-money effect. For high levels of the value-for-money rating, the model predicts a reverse relationship—given the concavity of the continuation profits, higher effort reduces the marginal benefits from strategic pricing and should thus be associated with higher prices.

To gain insights on the relationship between the continuation profit and the ratings, we regress the monthly total revenue of listings on a third degree polynomial of the overall star-rating, listing fixed effects, month fixed effects, and the number of days available

$$R_{it} = \beta_0 + \beta_1 \times rat_{it}^{ov} + \beta_2 \times (rat_{it}^{ov})^2 + \beta_3 \times (rat_{it}^{ov})^3 + \lambda_i + \lambda_t + av_{it} + \epsilon_{it}\,. \quad (18)$$

Since the continuation value is the discounted sum of monthly profits, the parameters of the third degree polynomial in Equation (18) should provide a good approximation of the relationship between the star-rating and the continuation value. A third degree polynomial is necessary to allow for the possibility of convexity and concavity over the domain of a non-decreasing function. The estimated relationship between ratings and revenues based on the estimated parameters of the rating polynomial is shown in Figure 3. The results are in line with a convex relationship in the lower range of the domain and a concave relationship in the upper range.[17]

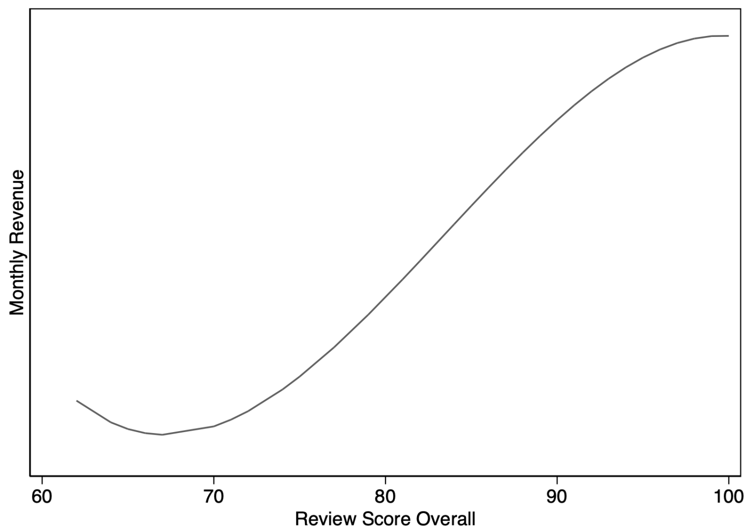


Figure 3: Relationship between monthly revenues and overall rating

.

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

[17]The choice of a third order polynomial in Equation (18) can be further justified when analyzing the relationship non-parametrically: To do so, we first obtain the residuals from a regression of the total monthly revenue on listing fixed effects, month fixed effects, and the days available. We then estimate the relationship between these residuals and the overall star-rating using a local-linear regression. Figure 6 of Appendix C reports the corresponding results. Additionally, Figure 7 reports the results for the non-parametric estimation of the relationship between the revenue and the value-for-money rating. The results support the choice of a third order polynomial in Equation (18).

lower levels of the rating distribution, and a positive correlation in the upper range.

To assess the model prediction, we create two groups characterized by the median aggregate star rating of a listing. The first bin captures all listings with a median star rating strictly lower than five, the second bin captures all listings with a median star rating equal to five. Define by b_k the kth bin defined by the five star threshold. We estimate the following regression

$$eff_{it} = \sum_{k=1}^{k=2} \beta_j log(p_{it}) I\{\tilde{rat}_i^{ov} \in b_k\} + X'_{it}\gamma + \lambda_i + \lambda_t + \epsilon_{it} \, . \tag{19}$$

$\tilde{rat}_i^{ov}$ denotes the the median rating of listing i and I is the indicator function which takes the value one if the median lies in the respective bin. As usual, we control for listing and month fixed effects. Our dependent variables are the three rating categories associated with effort (cleanliness, communication, and check-in) and the two available direct effort measures (host response rate and host response time). The X matrix contains the second order polynomial of the number of reviews to account for the averaging of the rating measures.

Table 4: Price-Effort Regression Results

	(1) Cleanliness	(2) Communication	(3) Checkin	(4) Response rate	(5) Response time
Price (log) - Rating < 5	-0.14	-0.26**	-0.17*	-0.00	-0.01
	[-0.34,0.06]	[-0.42,-0.10]	[-0.34,-0.01]	[-0.01,0.00]	[-0.03,0.01]
Price (log) - Rating = 5	0.18	-0.07	-0.04	-0.00	-0.03**
	[-0.06,0.42]	[-0.27,0.12]	[-0.24,0.15]	[-0.01,0.01]	[-0.06,-0.01]
Constant	93.62***	98.21***	97.60***	4.57***	0.72***
	[92.93,94.32]	[97.65,98.76]	[97.03,98.16]	[4.56,4.59]	[0.65,0.79]
Observations	182498	182395	182294	183099	183893

Notes: *, **, *** indicate statistical significance at the five, one, and 0.1 percent level, respectively. The square brackets show 95 percent confidence intervals.

The results from Table 4 are mixed. For the effort ratings, we observe a pattern that is consistent with decreasing complementarity. The same is not true with the direct effort measures we observe. We note that the response rate and the response time might only capture partial aspects of the overall effort hosts dedicate to their guests. The effort ratings might provide a more complete measure by also incorporating aspects not captured by the direct measures such as friendliness or level of detail of information.

5 Price-Rating Dynamics

In this section, we turn to a dynamic analysis of the relationship between prices and ratings. Hypothesis 4 proposes that sophisticated hosts offer relative discounts when entering the platform. Hosts charging low prices when entering the market will obtain better ratings than hosts that enter the market charging higher prices. As a result, hosts with low entry prices will be able to charge higher prices once ratings have stabilized, i.e., after they have accumulated a sufficient number of ratings, which are higher because of the low entry prices.

We do not have a clear measure to differentiate between sophisticated and unsophisticated hosts. Instead, we analyse whether listings that charge a relatively lower price when entering the platform can benefit from it in later periods. Such a result would indicate that some hosts are indeed using strategic pricing and would be consistent with Hypothesis 4.

To analyze the impact of low entry prices, we estimate the following equation

$$y_{it} = \beta_0 + \sum_{j=1}^{j=6} \beta_t \times fp_i \times \mathcal{I}(t=j) + \alpha' X_{it} + \epsilon_{it}\,. \tag{20}$$

y_{it} denotes different outcome variables for listing i at time t. The outcome variables include the price, the number of bookings, different rating categories, or revenues. fp_i is a measure of the initial discount that listing i offers in the first period it enters the market.[18] The computation of fp_i is explained in the next paragraph. 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.[19]

To measure the initial discount that a host offers, we compare the price charged in the very first month to the average price charged in the following months. Formally, we calculate:

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

[19]The location fixed effects are based on statistical, geographical units called IRIS. They were defined by the INSEE to capture areas with similar population sizes. We have 990 different IRIS in our sample.

$$fp_i = 1 - \frac{p_{i1}}{\bar{p}_{i2-6}} . \tag{21}$$

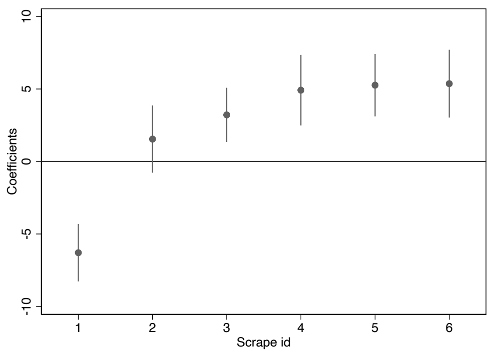
A negative value of fp_i arises if the initial price charged is higher than the average charged in subsequent periods. A positive value captures an initial discount. Note that the maximum discount is naturally one. β_t in Equation (20) can be interpreted as the effect of a 100 percent first-period discount on the outcome in period t. To aid the interpretation of the following results, we report the estimates scaled by the median discount (8.5%).

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

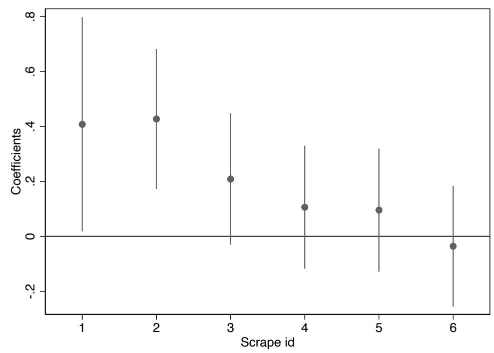
Figure 4 shows the estimated β_t-coefficients for different outcome variables. Figure 4a shows that listings with a larger initial discount tend to charge lower prices in the first period than those that do not offer an initial discount. This difference vanishes in the second period and reverses in the subsequent periods. Note that the shape of this curve is to some extent predetermined by the way we construct the discount variable. However, the level of the curve is not clear ex ante. For example, a curve lying below zero throughout would indicate that only listings that charge low prices in the "long run" offer an initial discount.

Figure 4b suggests that the initial discount seems to pay off in terms of the value-for-money rating. In the first months, those listings that offer an initial discount receive on average better value-for-money ratings. This is consistent with the dominant value-for-money effect documented in Section 4.1. When the prices increase in the subsequent periods, this advantage vanishes. Note, however, that while prices in the later periods are on average higher than those charged by listings without an initial discount, this does not translate into a penalty in terms of the value-for-money rating. The pattern looks broadly similar (albeit estimated with lower statistical precision) for most other rating categories, except for the location rating. This is intuitive, as we would expect the selection effect to play a larger role for the location rating. We report these results in Appendix D.

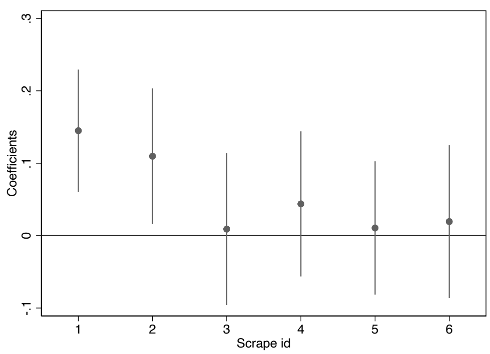
The results are similar when analysing the number of bookings. Figure 4c shows that the initial discount seems to draw additional bookings in the first



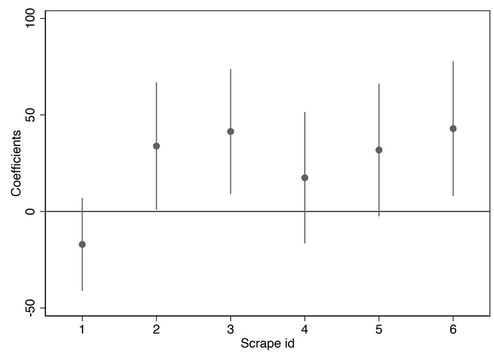
(a) y_{it} = Transaction prices



(b) y_{it} = Value-for-money rating



(c) y_{it} = Number of bookings



(d) y_{it} = Revenues

Figure 4: β_t for different y_{it} (scaled at median discount of 8.5%)

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 number of bookings would suggest, listings that set an initial discount are able to generate higher revenues in subsequent periods (see Figure 4d).

6 Conclusion

We investigate whether hosts on the short-term accommodation platform Airbnb can influence their ratings through strategic price setting. In a simple theoretical framework, we suggest that a higher price has two opposing effects on ratings. First, higher prices result in lower ratings due to a lower value for money. Second, higher prices result in higher ratings due to a self-selection of travelers into booking: only travelers with a high preference for the listing will book it. The net effect of prices on ratings depends on which of these effects, the value-for-money effect or the selection effect, dominates.

Using data on Airbnb transactions and corresponding ratings in Paris in 2017, we find that higher prices reduce most rating categories, suggesting that the value-for-money effect dominates the selection effect. The relationship is most prominent for the value-for-money rating. However, we do not find such a relationship for the location rating. We argue that this result is in line with a selection effect which should arguably be more important for the location rating, due to 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 predicts that hosts can use effort as another strategic control variable to affect their ratings. Whether price and effort are strategic complements or substitutes depends on whether future profits are convex or concave in the rating.

To assess the curvature of the continuation value empirically, we regress listing revenues on a third-degree polynomial of the overall rating as well as other control variables. Our results suggest that the continuation value is convex in ratings for lower ratings and becomes concave when ratings are high. Given this result, our theoretical model predicts that lowly-rated hosts should use effort as a strategic complement of pricing (i.e. exert more effort) whereas highly-rated hosts should use effort as a strategic substitute of pricing. Consistent with this prediction, we find suggestive evidence that hosts of lower-rated listings exert more effort.

Given these insights, we would expect that hosts can benefit from pricing strategically when entering the platform. In an analysis of entry pricing, we find results that are in line with this expectation: Listings with a relatively lower price when entering receive better value-for-money ratings and more bookings early on, allowing them to charge higher prices and realise higher revenues in subsequent periods.

Our paper provides a framework that reconciles prior results on the relationship between prices and ratings. Our empirical results suggest that, indeed, both a value-for-money as well as a selection effect seem to be at play, with the value-for-money effect being dominant for ratings on Airbnb. Furthermore, we show that hosts can affect their ratings through their price-setting and effort.

Future research could focus on whether such strategic pricing affects the informativeness of the rating system. In our theoretical model, ratings enter utility in a reduced form. However, if strategic pricing affects ratings, these ratings could become less informative of the true quality of the listing as a result. Such insights would also 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.

Fradkin, A., Grewal, E., and Holtz, D. (2020). Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb.

He, S., Hollenbeck, B., and Proserpio, D. (2021). The Market for Fake Reviews. Working Paper.

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

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*, forthcoming:1–19.

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.

Stenzel, A., Wolf, C., and Schmidt, P. (2020). Pricing for the Stars: Dynamic Pricing in the Presence of Rating Systems. 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.

Appendix

A Derivations

A.1 Strategic Seller Problem

Consider the first-order conditions from the Bellman equation

$$1 + \mu(\Psi_t^v) + \nu(\Psi_t^e) - 2p_t \tag{22}$$

$$+ \delta \left(\frac{d}{d\Psi^v} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^v}{dp_t} + \frac{d}{d\Psi^e} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^e}{dp_t} \right) = 0 \tag{23}$$

$$-ce_t + \delta \left(\frac{d}{d\Psi^v} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^v}{de_t} + \frac{d}{d\Psi^e} V_{t+1}(\Psi_{t+1}^v, \Psi_{t+1}^e) \frac{d\Psi_{t+1}^e}{de_t} \right) = 0. \tag{24}$$

Clearly, the value is increasing in the state variables—the ratings Ψ_t^v and Ψ_t^e—as 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 rating generation that

$$\frac{d}{dp_t} \Psi_{t+1}^j = \frac{1}{t+1} \frac{d}{dp_t} \psi^j \tag{25}$$

$$\frac{d}{de_t} \Psi_{t+1}^j = \frac{1}{t+1} \frac{d}{de_t} \psi^j. \tag{26}$$

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

$$\begin{aligned} \frac{d}{dp_t} \psi^v &= \frac{d}{dp_t} \varphi^v(\theta, e, \omega^e(p), p) \\ &= \varphi_\omega^v \frac{d\omega^e}{dp_t} + \varphi_p^v \\ &= \varphi_\omega^v \frac{1}{2} + \varphi_p^v \end{aligned} \tag{27}$$

implying that the overall effect of the price on the value for money rating can be either positive or negative. It is negative whenever $\frac{-\varphi_p^v}{\varphi_\omega^v} > \frac{1}{2}$., i.e., 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 effort according to

$$\frac{d}{de_t}\psi^v = \frac{d}{de_t}\varphi^v(\theta, e, \omega^e(p), p) = \varphi^v_e > 0, \tag{28}$$

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

$$\frac{d}{dp_t}\psi^e = \frac{d}{dp_t}\varphi^e(\theta, e, \omega^e(p), p) = \varphi^e_p \leq 0 \tag{29}$$

$$\frac{d}{de_t}\psi^e = \frac{d}{de_t}\varphi^e(\theta, e, \omega^e(p), p) = \varphi^e_e > 0. \tag{30}$$

We can rewrite the first-order conditions as

$$1 + \mu(\Psi^v_t) + \nu(\Psi^e_t) - 2p_t + \frac{\delta}{t+1}\left(\frac{d}{d\Psi^v}V_{t+1}\left(\varphi^v_\omega \frac{1}{2} + \varphi^v_p\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi^e_p\right) = 0 \tag{31}$$

$$-ce_t + \frac{\delta}{t+1}\left(\frac{d}{d\Psi^v}V_{t+1}\varphi^v_e + \frac{d}{d\Psi^e}V_{t+1}\varphi^e_e\right) = 0. \tag{32}$$

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

$$e_t = \frac{\delta}{c(t+1)}\left(\frac{d}{d\Psi^v}V_{t+1}\varphi^v_e + \frac{d}{d\Psi^e}V_{t+1}\varphi^e_e\right) > 0. \tag{33}$$

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

$$p_t = \underbrace{\frac{1 + \mu(\Psi^v_t) + \nu(\Psi^e_t)}{2}}_{p^m_t} + \frac{\delta}{2(t+1)}\left(\frac{d}{d\Psi^v}V_{t+1}\left(\varphi^v_\omega \frac{1}{2} + \varphi^v_p\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi^e_p\right). \tag{34}$$

Inspecting (34), 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

$$\frac{d}{d\Psi^v}V_{t+1}\left(\varphi^v_\omega \frac{1}{2} + \varphi^v_p\right) + \frac{d}{d\Psi^e}V_{t+1}\varphi^e_p. \tag{35}$$

In particular, prices are lower than the myopic prices whenever

$$-\frac{\varphi^v_\omega \frac{1}{2} + \varphi^v_p}{\varphi^e_p} > \frac{\frac{d}{d\Psi^e}V_{t+1}}{\frac{d}{d\Psi^v}V_{t+1}}. \tag{36}$$

Moreover, for the special case where the price does not affect the effort rating, i.e. when $\varphi_p^e = 0$, the sign of (35) is fully determined by the sign of $\varphi_\omega^v \frac{1}{2} + \varphi_p^v$, so that

$$p_t < p_t^m \iff \frac{-\varphi_p^v}{\varphi_\omega^v} > \frac{1}{2}. \tag{37}$$

This is the same condition as in (27), which is intuitive—as prices only affect 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, i.e., if and only if the direct price effect dominates the indirect selection effect.

A.2 Strategic pricing and effort

To analyze this, 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 (31) and (32).

$$\begin{aligned}
\frac{dp_t}{de_t} &= \frac{\delta}{2}\frac{d}{de_t}\frac{dV_{t+1}}{dp_t} \\
&= \frac{\delta}{2}\frac{d^2V_{t+1}}{de_t dp_t} &(38)\\
&= \frac{\delta}{2(t+1)^2} &(39)\\
&\left(\frac{d^2V_{t+1}}{d(\Psi^v)^2}\frac{d\psi^v}{dp_t}\frac{d\psi^v}{de_t} + \frac{d^2V_{t+1}}{d(\Psi^e)^2}\frac{d\psi^e}{dp_t}\frac{d\psi^e}{de_t} + 2\frac{d^2V_{t+1}}{d\Psi^v d\Psi^e}\left(\frac{d\psi^e}{dp_t}\frac{d\psi^v}{de_t} + \frac{d\psi^e}{de_t}\frac{d\psi^v}{dp_t}\right)\right)\\
&+ \frac{\delta}{2(t+1)^2}\left(\frac{dV_{t+1}}{d\Psi^v}\frac{d^2\psi^v}{dp_t de_t} + \frac{dV_{t+1}}{d\Psi^e}\frac{d^2\psi^e}{dp_t de_t}\right). &(40)
\end{aligned}$$

Whether effort and strategic price adjustments are substitutes or complements depends both on the sign of (40) and the sign of (35)—we say that strategic price management and effort are complements if an increase in effort leads to a further price adjustment in the direction which increases future profits at the expense of flow profits. To better understand the economic forces at play, suppose that $\varphi_p^e = 0$ and $\varphi_{pe}^v = 0$, i.e. that the price does not affect the effort rating, and that there are no cross-effects between price and effort in the value-for-money rating. This allows us to write

$$\frac{dp_t}{de_t} = \frac{\delta}{2(t+1)^2}\frac{d\psi^e}{de_t}\frac{d\psi^v}{dp_t}\left(\frac{d^2V_{t+1}}{d(\Psi^v)^2} + 2\frac{d^2V_{t+1}}{d\Psi^v d\Psi^e}\right). \tag{41}$$

Note that the sign of the term $\frac{d\psi^e}{de_t}\frac{d\psi^v}{dp_t}$ is fully determined by $\frac{d\psi^v}{dp_t}$, which simultaneously determines whether higher or lower prices induce higher ratings. Thus, price and effort are complements whenever the sign of $\frac{dp_t}{de_t}$ is the same as the sign of $\frac{d\psi^e}{de_t}\frac{d\psi^v}{dp_t}$. If this is the case, a higher effort incentivizes a price change in the direction that further increases ratings. Thus, whether effort and price are substitutes or complements is determined by the sign of the term

$$\frac{d^2V_{t+1}}{d(\Psi^v)^2} + 2\frac{d^2V_{t+1}}{d\Psi^v d\Psi^e}, \tag{42}$$

which measures the curvature of the continuation profits with respect to changes in the rating. If this term is positive, an increase in effort increases the marginal benefit of increasing ratings further—price and effort are complements. If it is negative they are substitutes.

B Impact of Ratings on Host Revenues

The model introduced in the main text assumes that ratings positively affects profits. Otherwise, there would be no reason for sellers to try to control ratings. In this section, we empirically validate this model assumption.

There is by now an extensive literature on the effect of ratings on seller performance. Luca (2016), for example, exploits a rounding threshold on the Yelp platform to study the impact of ratings on revenues for a sample of restaurants. The analysis we present in this Appendix follows a similar logic.

From our data, we are able to observe a granular overall quality measure ranging on a scale from 20 to 100 in increments of one unit. This granular scale is not observed by travelers. At the time our data were sampled, travelers only observed a less granular quality measure that ranged from 1 to 5 stars in increments of half a star.

The granular quality measure determines the number of stars shown to travelers. A granular quality measure in the interval $[20, 25)$ corresponds to one star, $[25, 35)$ to one and a half stars, $[35, 45)$ to two stars. This relationship continues until the last interval, from $[95, 100]$, which corresponds to five stars.

The granular quality measure lends itself well to study the impact of the overall rating on revenues in an regression-discontinuity-design framework. The data allow us to compare listings which are almost identical with respect to the overall non-salient rating but which differ with respect to the salient rating.

For example, listings with an overall rating of 94 and 95 are almost identical with respect to the overall non-salient rating measure but differ saliently in the overall star rating shown to travelers. To analyze the impact of salient rating thresholds on performance, we run the following regression:

$$y_{it} = \beta_0 + \beta_1 \mathcal{I}_{bw}(rat_{it} > \tau^\star) + X_{it} + \mu_i + \lambda_t + \epsilon_{it} \tag{43}$$

y_{it} denotes the revenue, μ_i and λ_t listing and time fixed effects, respectively. X_{it} contains time-varying control variables which are potentially important to explain revenues, such as the number of days a listings is available for booking in a month and the number of reviews it accumulate.

In Equation (43), bw denotes the bandwidth chosen for the regression discontinuity design and $\tau^\star$ denotes the threshold. $\mathcal{I}_{bw}(rat_{it} > \tau^\star)$ is an indicator variable that takes the value one if the non-salient rating exceeds the threshold. For $bw = 0.5$ and $\tau^\star = 80.5$, β_1 captures the difference in y_{it} between listings with a rating of 81 and 80.

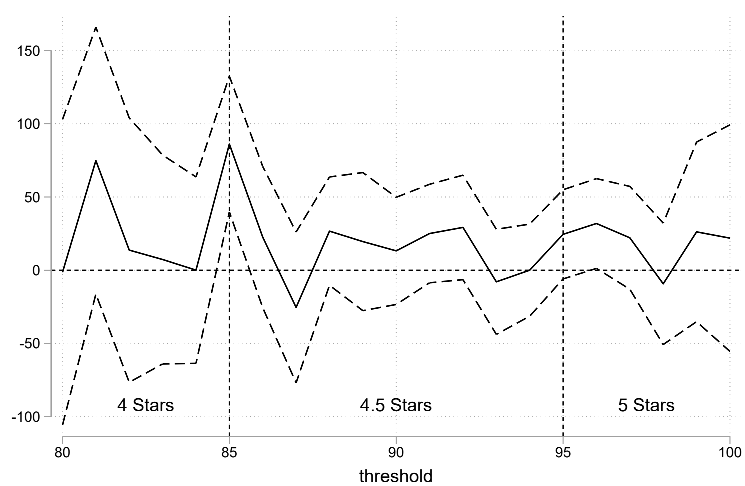


Figure 5: Impact of incremental granular rating change on revenues

In the following, we fix $bw = 0.5$ and let $\tau^\star = \{80.5, 81.5, 82.5, ..., 99.5\}$. Thus, we estimate the coefficient $\mathcal{I}$ for listings which are adjacent in the non-salient overall rating. We run a separate regression for each threshold.

For the threshold values of $\tau^\star = \{84.5, 94.5\}$ the difference in the non-salient measure will result in changes in the salient rating measure, otherwise not. The first threshold marks the transition from 4 to 4.5 stars and the second threshold the transition from 4.5 to 5 stars. We focus on these two thresholds in the upper range of the rating distribution because we lack observations for a meaningful analysis in the lower range.

Figure 5 shows the estimator for β_1 in Equation (43) for the different values of the threshold. There is a clear discontinuity at the salient threshold from 4 to 4.5 stars in the order of magnitude of 100 euro per month. Non-salient changes are generally associated with a positive but insignificant effect on revenues. A salient threshold change from 4.5 to 5 stars appears to have no impact on revenues.

Note that the significant effect at the first threshold and the insignificant effect at the second threshold are consistent with our stylized finding of a convex-concave continuation value function from ratings: Marginally increasing ratings at a lower baseline level increases revenues more than a marginal increase in ratings at higher levels. The results of our analysis are consitent with a causal link from ratings to revenues. Host have an incentive to attempt to influence ratings by strateic price setting.

C Non-Parametric Estimation of Continuation Value

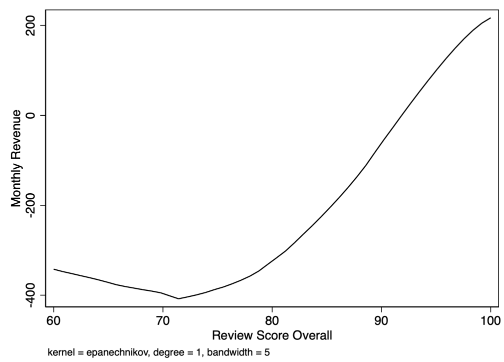


Figure 6: Relationship between monthly revenues and overall rating
Local Polynomial Regression

.

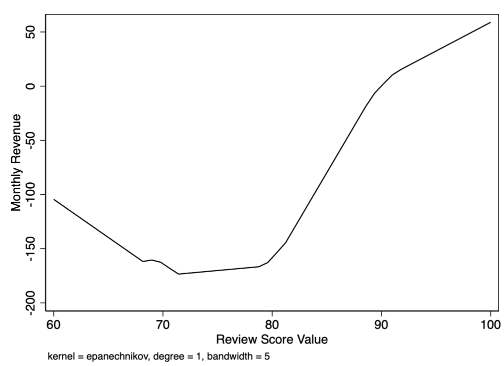
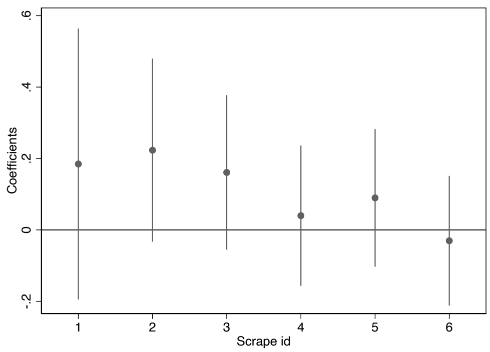


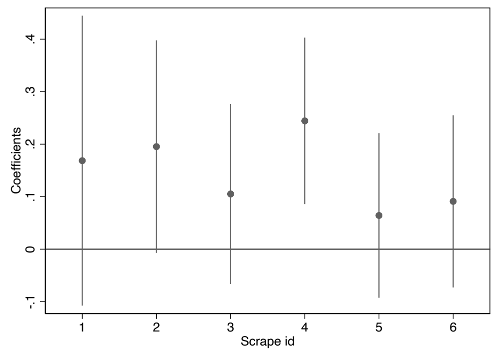
Figure 7: Relationship between monthly revenues and value rating
Local Polynomial Regression

.

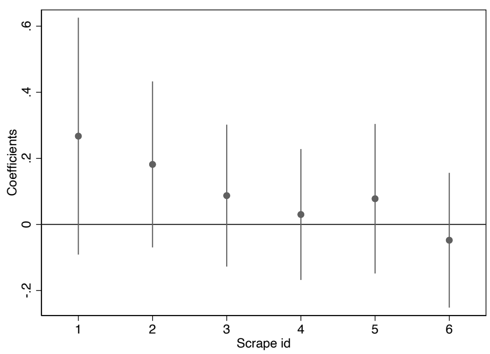
D Other Price-Rating Dynamics



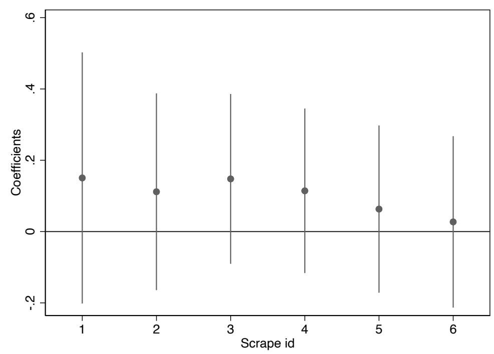
(a) y_{it} = Overall rating



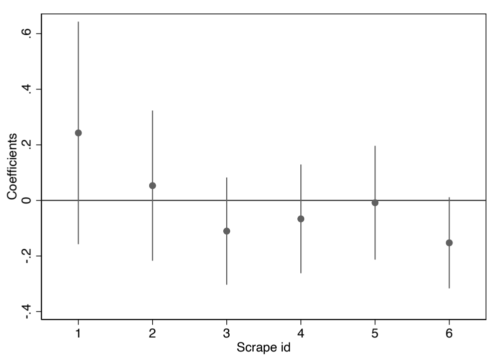
(b) y_{it} = Location rating



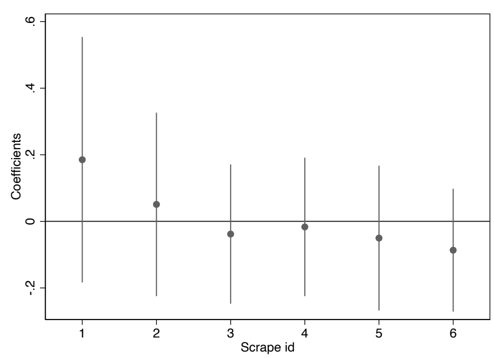
(c) y_{it} = Accuracy rating



(d) y_{it} = Cleanliness rating



(e) y_{it} = Communication rating



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

Figure 8: β_t for different y_{it} (scaled at median discount of 8.5%). Corresponds to analysis described in Section 5.