How to “buy” honest reviews. Experimental evidence of the impact of prices on online reviews*

Przemyslaw Jeziorski[†] Konstantina Michelidaki[‡]

June 17, 2022

Abstract

The paper documents the causal impact of transaction price on the subsequent product review submitted by the customer. The evidence is obtained using a Large Scale Randomized Field Experiment which introduced a random component into over 3 million nightly rates of short-term rentals. Using this variation, we show that exogenous 25% price decrease increases subsequent star rating by nearly one-fourth of standard deviation. Conversely, we also demonstrate that an exogenous increase of the most recent review by a standard deviation increases revenue from a subsequent transaction by at least 25%. These relationships imply that firms may *invest* in better reviews by lowering their price in addition to investing directly in product quality; conversely, firms may *monetize* good reviews by raising prices. We discuss implication of these results for optimal pricing and market efficiency.

*We would like to thank Andrey Fradkin, Ganesh Iyer, and Miguel Villas-Boas for helpful discussion and suggestions. Both authors are co-founders and hold equity in Keybee Inc., the data supplier for this paper. Both authors received income from Keybee Inc. in excess of $10,000, which was unrelated to the research in this paper. Both authors and are on the board of directors of Keybee Inc. Konstantina Michelidaki is a CEO of Keybee Inc.

[†]przemekj@berkeley.edu, UC Berkeley, Haas School of Business

[‡]nadia@keybeehosting.com, Keybee Inc.

1 Introduction

Reputation systems based on customer reviews emerged as a prominent feature of modern online marketplaces. The main purpose of the reviews is to increase market efficiency by lowering asymmetric information about product quality between buyers and sellers.[1] Indeed, numerous findings suggest that customer reviews are an informative signal about product quality.[2] However, the literature identified several mechanisms impeding the information value of the reviews[3], including self-selection, review externalities, and fake reviews. We show firms can "buy" favorable organic reviews by temporarily lowering their prices.

To obtain the causal impact of prices on reviews, we used a naturally occurring field experiment in which an optimal pricing algorithm randomly perturbed over 3 million posted prices on a leading online platform offering short-term rentals.[4]. We show that an exogenous 25% decrease in transaction price leads to an increase in review rating by approximately a fourth of standard deviation. We also demonstrate that the effect is more prominent for accommodations with fewer photos and shorter descriptions likely characterized by a greater degree of asymmetric information.

Conversely, we find a positive causal relationship between the most recent review and subsequent transaction price. In particular, we show that increase in the star rating of the most recent sale increases price of a subsequent transaction by at least 25%.[5] Our causal interpretation of the link between reviews and future performance relies only on experimental variation. In contrast, the literature on the impact of reviews typically relies on non-random natural experiments (ex. regression discontinuity or difference-in-differences)

[1]Classical reference showing detrimental role of asymmetric information is Akerlof (1978). Beyond adverse selection, in the empirical setting considered in this paper (short-term rentals), asymmetric information may also lead to more nuanced frictions such as statistical discrimination, see Laouénan and Rathelot (2022).

[2]See Chevalier and Mayzlin (2006), Cabral and Hortacsu (2010), and Anderson and Magruder (2012).

[3]Lewis (2011) shows that online market places are subject to significant residual asymmetric information, despite the ubiquity of reviews.

[4]The prices were perturbed by multiplicative shock was distributed uniformly on the interval [0.9,1.1] and was applied independently (IID) across *pricing events*. A pricing event combines a property, the date the property is occupied, and a pricing day. New shocks are applied daily between midnight and 1 am local time to random 30% of nightly rates of future unrented nights within the next 2-year window. The dataset consists of 341 unique rental properties and 1,003 rental days, repriced daily. Accounting for customer entry and attrition, we obtain over 3 million pricing events that obtained a non-trivial random shock.

[5]The dynamic pricing algorithm drops the price as the check-in date approaches to obtain over 90% capacity utilization; thus, transaction prices, not conversion rates, are the most informative measure of profitability in our setting.

because it is difficult to introduce truly random variation into the reviews. To obtain such random variation we use a convenient structure of our field experiment; that is, our price shocks are identically and independently distributed (IID) across accommodations and over time. In such a case, shocks to lagged transaction prices are suitable instruments for lagged reviews since (i) they are uncorrelated with unobserved determinants of the current price, and (ii) they are correlated with lagged reviews as discussed in the previous paragraph.

The impact of transaction price on subsequent reviews, and the converse demand feedback loop have important implications for pricing strategy. Immediately, since the review score is persistent, optimal pricing becomes dynamic, even if other market features would only require static pricing[6]. Basic pricing theory prescribes that the firm should change its price, if contemporaneous demand elasticity changes due to a change in review rating.[7] Beyond this mechanism, our results suggest further incentives for the firm to change its price in response to review ratings. In particular, a firm may lower its price to *invest* in better reviews, taking into account the ability *monetize* good reviews in the future. This mechanism has market design and regulatory implications since inferior firms with deep pockets may be able to generate market dominance via product reviews.[8]

The data suggest several possible mechanisms that could generate a relationship between transaction prices and subsequent customer satisfaction voiced in the review. One possibility is *moral hazard* caused by a lower consumer surplus of infra-marginal buyers. If buyers incorporate surplus into their ratings, higher prices will lower reviews, even under perfect information about quality. In our empirical setting, the platform aims to directly capture consumer surplus by soliciting *value* rating.[9] We find that the impact of price increase on the value rating is of similar magnitude as on the overall rating, suggesting that lower consumer surplus may be a driving force behind our results.

The second possible mechanism is more nuanced and results from an advantageous selec-

[6]See Stenzel, Wolf, and Schmidt (2020) for theoretical analysis.

[7]For example, Zhong (2021) shows that sellers raise prices after a discontinuous jump in their star rating.

[8]The mechanism is similar to *predatory pricing* and *dominance* in the markets with learning-by-doing, see Cabral and Riordan (1994) Cabral and Riordan (1997); or building *brand equity* via costly advertising and promotion, see Dubé, Hitsch, and Manchanda (2005) and Borkovsky, Goldfarb, Haviv, and Moorthy (2017).

[9]Airbnb defines value as “did the guest feel that the listing provided good value for the price?”.

tion of buyers as the posted price increases.[10] In particular, a price hike removes marginal buyers who have the lowest preference for the product. Consequently, depending on the curvature of customers' utility function and the distribution of preference heterogeneity, the price hike may lead to either a decrease or increase in satisfaction of buyers and resulting reviews. The direction of the effect depends on the relative importance of moral hazard and advantageous selection.[11]

The third possible mechanism relies on the existence of asymmetric information. The prominence of customer reviews in the short-term rental market suggests a significant amount of uncertainty about product quality before purchase. A considerable degree of private information is not surprising since short-term rental accommodations are prone to larger quality dispersion than hotels. An online listing may have limited ability to communicate the idiosyncratic quality information; thus, the review system becomes a central revelation mechanism in most short-term rental platforms. Apart from the content of an online listing and its reviews, the posted price may be an additional signal of quality, as in Wolinsky (1983). We demonstrate that, in theory, asymmetric information is likely to exacerbate the negative impact of price on reviews because after the price hike, consumers generally expect higher quality.[12] The data reveals that even the ratings related to objective measures, such as *location* or *check-in instructions* decrease significantly after the price hike. If consumers infer the quality of location or check-in instructions from posted price, a price hike may lead to ex-post regret.

Absent the experiment; the relationship between price and reviews is difficult to identify because of price endogeneity. In particular, prices and reviews are likely to be simultaneously determined by unobserved product characteristics. One can partially address this

[10]Pioneering work theoretical work demonstrate that selection of reviewers matters was done by Li and Hitt (2008) and Hu, Pavlou, and Zhang (2017).

[11]An anecdotal example of advantageous selection from the academic profession is instructor ratings. Some faculty believe that increasing the class difficulty (the price of taking the class) may result in better instructor reviews. The presumed increase in ratings may occur because higher difficulty encourages marginal students to drop the course. However, by a similar principle, increasing the difficulty without additional academic benefits would upset students who decided to stay in the class. The overall effect direction depends on the relative strength of selection and impact on infra-marginal students.

[12]Extending the student teaching rating example from Footnote 11, one could imagine that increasing perceived quality of the class may result in students seeking an academic gratification. However, if the class merely becomes harder without delivering additional gratification, the reviews should be more prone to decrease than in the model with complete information.

endogeneity by using panel data, leveraging within-product variation of transaction prices and reviews. Nevertheless, contemporary price and reviews are likely to be simultaneously determined by serially correlated shocks to unobserved quality, which may be caused by seasonality and any other time-varying demand shifters. As a result, the estimates based on panel data may still suffer from considerable bias. Moreover, the endogeneity in the panel data is particularly concerning for short-term rentals and other products price using automated dynamic pricing algorithms. Such algorithms automatically decrease the posted price of unsold check-in dates as the potential reservation date approaches. This decrease generates a significant degree of correlation between demand shifters and prices, even if the demand shifters are unobserved to the firm ex-ante.[13] This paper addresses price endogeneity with experimental data on exogenous random shocks to the price delivered by an exploratory pricing module employed by the firm supplying the data for this project. The paper considers a series of estimators demonstrating price endogeneity when relying only on observable listing characteristics. Finally, the paper presents causal instrumental variables estimator that leverages only experimental price variation.[14]

The paper is organized as follows. Section 2 contains an Industry Description. Section 3 discusses the data sample and experimental design. Model-free and model-aided results are contained in Section 4. The evidence of price cycles is contained in Section 5. Section 6 concludes.

2 Short-term rental industry

This paper focuses on the short-term rental (STR) industry.[15] STR is a business model in which landlords (henceforth called Hosts) rent furnished properties for a short time, compared to long-term rentals – primarily unfurnished properties rented on yearly contracts. The

[13]Another, more technical reason, for bias in panel-data estimators, is the usual issues present in a dynamic panel with fixed effects as in Arellano and Bond (1991). The bias occurs if the regressors are predetermined (not strictly exogenous). Unfortunately, past prices are likely to violate strict exogeneity if they correlate with contemporary or lagged demand shocks.

[14]The results in the paper are consistent with the pioneering work of Luca and Reshef (2021), who find that increases in prices lead to lower restaurant ratings on Yelp. Luca and Reshef (2021) uses observational data and applies item-week fixed effects.

[15]Another synonymous name for short-term rentals is vacation rentals.

global market valuation of the STR industry was estimated at $78 billion in 2021.[16] Depending on the geographical market, STR landlords must obtain permits to operate and abide by various special regulations. For example, many jurisdictions define short-term rental as any rental below 30 days. Tenants of STRs (henceforth called Guests) typically do not obtain tenant rights and must instead abide by tailored STR contracts. The closest competitors to STRs are hotels.

STRs have a long history and were typically executed via informal markets. Today, most short-term rentals are executed via organized online platforms like Airbnb, VBRO, Expedia, and Booking.com. These platforms facilitate the interaction between Hosts and Guests, from search, booking, and payment, to check-out and feedback. STR platforms provide a contractual framework and arbitrate disputes. A significant difference between hotels and STRs is the degree of ownership concentration. Most Hosts are individuals offering one or only a few properties to let.[17] In comparison, a hotel typically offers multiple rooms and could be a part of a larger chain. For this reason, the degree of asymmetric information about product quality has the potential to be more significant for a short-term rental than for the hotel. As a result, platforms have introduced elaborate review systems to mitigate the asymmetric information problem. The actual design of these systems varies. For example, Booking.com allows to rate properties but does not allow rating Guests. In comparison, VRBO and Airbnb allow for rating both but adopt double-blind review systems to avoid reciprocation and retaliation.[18] Possibly due to extensive asymmetric information, reviews became the leading driver of the demand for a particular STR. The reviews have the potential to drive the probability of booking conditional on examining the online property listing and are a significant driver of the platform's ranking and search algorithm.[19]

Despite the indisputable importance of reviews in modern marketplaces, including the STR platform in this study, a scarce amount of research describes customers' deliberation

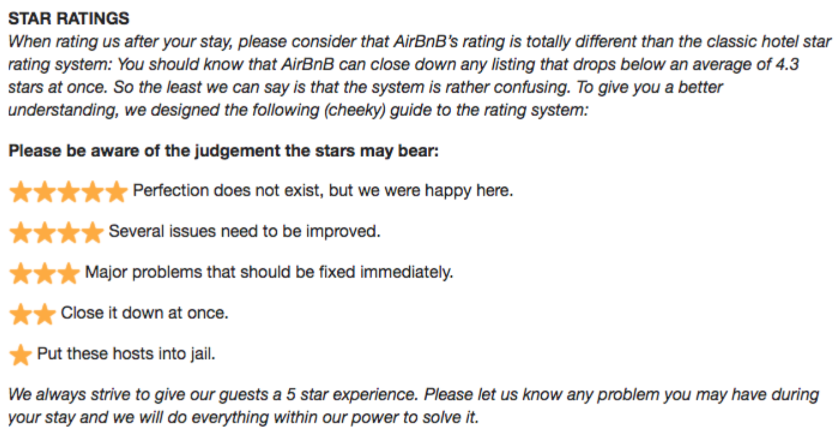
[16] Source "Vacation Rental Market Size, Share & Trends Analysis Report...," `https://www.grandviewresearch.com/industry-analysis/vacation-rental-market`, accessed 4/22/2022.

[17] According to Airbnb website accessed on 4/22/2022, the platform contains 6M listings and 4M+ Hosts, see `https://news.airbnb.com/about-us/`.

[18] See Fradkin, Grewal, and Holtz (2021)

[19] The authors of this paper obtained scraped data on online search rankings from one of the platforms and demonstrated a strong correlation between rating and search position even after controlling for property fixed effects.

process resulting in the review. Therefore, we postulate that overall rating may be driven by “value”, quality net of price, and raw quality experience, disregarding the price.[20] The STR platform solicits the feedback using several frames in which you rate customer experience using a star rating. Figure 1 contains the description of the star system communicated to the users by the STR platform.



STAR RATINGS

When rating us after your stay, please consider that AirBnB's rating is totally different than the classic hotel star rating system: You should know that AirBnB can close down any listing that drops below an average of 4.3 stars at once. So the least we can say is that the system is rather confusing. To give you a better understanding, we designed the following (cheeky) guide to the rating system:

Please be aware of the judgement the stars may bear:

★★★★★ Perfection does not exist, but we were happy here.

★★★★ Several issues need to be improved.

★★★ Major problems that should be fixed immediately.

★★ Close it down at once.

★ Put these hosts into jail.

We always strive to give our guests a 5 star experience. Please let us know any problem you may have during your stay and we will do everything within our power to solve it.

Figure 1: Description of Airbnb star rating. Source and copyright Airbnb Inc., fair use doctrine.

The most crucial frame displayed next to the listing name in the search result is the “overall rating.” The “overall” rating is an unaided frame, which simply asks the question “How was your stay at X’s place?” In addition to the unaided frame, the platform uses several aided frames that focus on distinct parts of the customer experience. Conveniently, from an empirical perspective, various review frames will enable us to understand customers’ deliberation process when leaving feedback.

Apart from the unaided review, Airbnb solicits reviews using 6 aided categories:

- Value
- Accuracy
- Check-in
- Communication

[20]Imagine booking an extremely cheap short-term rental for a work trip. After arrival, the location was not ideal, but you still rent it 5-stars because of the bargain deal. On the other hand, imagine paying extra for a sea-view room for your vacation in the middle of the season. Despite the exuberant price, you still rent the room 5-stars because of the magnificent sunset views.

- Location
- Cleanliness

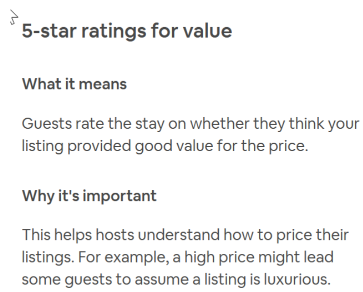


Figure 2: Screenshot from the Host Dashboard that explains the goal of the value rating solicitation. Source and copyright, Airbnb Inc., fair use doctrine.

The ratings in aided categories are not immediately visible to future consumers next to the listing name. Still, they are displayed to Guests on the "reviews" sub-page of the online listing. The aided frames aim to provide information to the Hosts to improve their accommodation along the specific dimensions. The latter 5 categories, apart from "value," aim to solicit various dimensions of accommodation quality. For example, the "location" category aims at improving "listing description" accuracy[21]. The former category, "value," has a different goal. According to the platform, "value" rating does not aim at quality improvement but at improving pricing. Figure 2 contains the description of the "value" frame that the platform communicates to the owners.

Looking at the "value" ratings, we would be able to estimate how important is consumer surplus during the feedback deliberation process. In the next Section, we present a stylized model of consumer preferences and deliberation that highlights the drivers of customer reviews.

3 Model

Despite the extensive literature on the impact of reviews on the level of demand, there is surprisingly little theoretical research on the customers' deliberation process when submitting

[21] According to the platform "Host Dashboard", accessed on 3/14/2022.

the review. For this reason, we developed a stylized theory model of customers' deliberation. The model is mathematically parsimonious by design because its primary goal is to formalize the subsequent discussion of the empirical relationship between prices and reviews. In particular, the model highlights possible mechanisms of review generation and postulates pathways via which transaction prices may affect customer satisfaction ratings, which are referenced when discussing empirical estimates.

Consider a continuum of customers and a single product. The product is a short-term rental accommodation for specific calendar dates at a specific property. Utility of the accommodation is given by

$$u = \alpha q - p,$$

where p is the price, and q is the vertical quality. Parameter α represents price-quality trade-off and varies across potential customers. Large values of α represent customers with a higher willingness to pay for quality. We assume that α is distributed with a CDF denoted by $F(\cdot)$.

The consumers may purchase the accommodation or outside option with utility $u_0 = 0$. By design, the utility model abstracts from many prominent features of the short-term rental market, such as horizontal differentiation[22] and competition. This decision allows us to highlight basic mechanisms linking transaction price with subsequent reviews without introducing extensive analytical complexity.

As explained in Section 2, customer reviews are solicited using various aided and unaided frames. Consider frame f with an expected review production function $R^f(u, q)$, which. In particular, we postulate that the expected review may depend on consumer surplus u, capturing the "value" component, and q, capturing dependence on raw quality. It is reasonable to expect that R^f is weakly increasing in both components. The relative importance of u and q depends on the review frame – a specific circumstance for generating feedback. This formulation deliberately ignores the multidimensional nature of product quality solicited by various review frames provided by the STR platform. However, varying degree of dependence on u and q enables us to isolate 3 distinct cases: (i) aided "value" frame, which puts

[22]Bondi and Stevens (2019) demonstrates that horizontal differentiation is an important factor driving reviews.

relatively more weight on u, (ii) aided quality frames, which put relatively more weight on q, and (iii) unaided "overall" quality frame that is potentially a more balanced function of u and q. We will drop superscript f to simplify the notation in the remainder of the paper. The context implies the frame.

First, consider a case in which quality q is known to the customer before purchase. Further, consider a frame f that does not depend on consumer surplus, u, but is entirely determined by quality, q. It is straightforward to notice that transaction price would not affect the frames' reviews. Thus, within our simple model, under complete information, the dependence of review on the transaction price would indicate that consumer surplus influences review deliberation. We expect this to occur in the "value" frame, but it may or may not occur in aided quality frames and the unaided frame.

There are two mechanisms via which review depends on the price. Consider a price increase from p to p'. Infra-marginal consumers with quality taste $\alpha > \frac{p'}{q}$ purchase the product before and after the price increase. If review frame R depends on u infra-marginal consumers would leave a lower review under price p'. We denote this effect as *moral hazard* because it results from the fixed consumer changing their review rating.

Second mechanism is *advantageous selection* of buyers. In most markets, including the STR platform, the review may only be left by consumers that purchased the product. A price increase selects buyers with a higher preference for quality, thus, truncating the distribution of α. Consumers with α, such that $\frac{p'}{q} > \alpha > \frac{p}{q}$, would have purchased with the lower price, p, but do not purchase with the higher price, p'. Consequently, an increase in price removes buyers with the lowest u, potentially increasing the reviews. The final impact of a price hike on reviews is a combination of advantageous selection and moral hazard and may increase or decrease reviews.

We examine moral hazard and advantageous selection mechanisms formally. To further simplify the argument we fix product quality to 1 and drop direct dependence of R on q. We also assume that R is differentiable in u. The average review is given by

$$\int_p^\infty R(\alpha - p)\frac{f(\alpha)}{1 - F(p)}d\alpha$$

Taking the derivative with respect to p we obtain that review drops as price increases, if and only if

$$-R(0)\frac{f(p)}{1-F(p)}+\int_p^{\infty}\left[R(\alpha-p)\frac{f(\alpha)f(p)}{(1-F(p))^2}-R'(\alpha-p)\frac{f(\alpha)}{1-F(p)}\right]d\alpha<0$$

After some manipulations we obtain

$$\int_p^{\infty}\left\{\big[R(\alpha-p)-R(0)\big]H(p)-R'(\alpha-p)\right\}f(\alpha)d\alpha<0,$$

where $H(p)$ is the hazard rate of the marginal consumer. First term inside the integral represents advantageous selection, after removing consumer marginal consumer. The larger the hazard rate $H(p)$ the more marginal consumers we remove by raising p and were allocate more mass to infra-marginal consumers who leave higher reviews, since $R(\alpha-p)-R(0)>0$. Second term represents moral hazard for each infra-marginal consumer, α. It suffices that moral hazard is larger than advantageous selection on average. If we impose the inequality point-wise, we obtain a simple ratio, linking the review-sensitivity of infra-marginal consumers to the hazard rate of marginal consumers

$$\frac{R'(\alpha-p)}{R(\alpha-p)}>H(p),\forall\alpha>p.$$

It is instructive to consider a special case for which $R'(\alpha-p)=0$, if $\alpha-p>\bar{u}$. In the case of STR accommodations, a flat review function may occur for luxury segment consumers. Such customers would review only using raw quality q, formally $R(\cdot,q)=R(q)$. We may see an advantageous selection in such a case since the increasing price would remove price-minded consumers and shift focus to the luxury segment that focuses on raw quality. The potential positive impact of prices on future reviews resulting from advantageous selection may incentivize Hosts to increase prices of luxurious properties beyond myopic optimum.

Next, we consider a case encompassing the asymmetric information about quality. The communication of the STR platform on Figure 2 indicates asymmetric information by stating that "...a high price might lead some Guests to assume that listing is luxurious." This statement presumes that (i) quality of accommodation is ex-ante unknown, and (ii) the market is in a (partially) separating equilibrium, in which posted price is a signal of quality. We consider a fully separating case, as in Wolinsky (1983) that generates a deterministic consumer belief quality schedule, $Q(p)$,[23] where p is the posted price. We postulate that $Q(p)$ is increasing.

In the following analysis, we abstract from general equilibrium effects by which quality schedule $Q(p)$ changes. Instead, we aim to highlight incentives for unilateral price deviations, as in our field experiment. We leave general equilibrium for further research. The schedule is not likely to change for unilateral price deviations since each Host is small and has little power to move market-level beliefs. Instead, unilateral price changes only generate movement along with the schedule.[24]

Again we set the true quality equal to $q = 1$ and consider a price increase from p to p'. We also assume that the true quality is revealed after the purchase so that the review is given by $R(\alpha q - p)$. Let $G(p) = \frac{p}{Q(p)}$ be the lowest α at which product is bought. The average review is given by

$$\int_{G(p)}^{\infty} R(\alpha - p)\frac{f(\alpha)}{1 - F(G(p))}d\alpha$$

Taking the derivative with respect to p we obtain that review drops as price increases, if and only if

$$-R(G(p) - \alpha)\frac{f(G(p))}{1 - F(G(p))}G'(p)+ \int_{G(p)}^{\infty} \left[R(\alpha - p)\frac{f(\alpha)f(G(p))G'(p)}{(1 - F(G(p)))^2} - R'(\alpha - p)\frac{f(\alpha)}{1 - F(G(p))}\right] d\alpha < 0$$

Suppose that the company deviates from equilibrium price p, such that $Q(p) = q = 1$, so that $G(p) = p$. We obtain

$$\int_{p}^{\infty} \left\{\big[R(\alpha - p) - R(0)\big]G'(p)H(p) - R'(\alpha - p)\right\}f(\alpha)d\alpha < 0,$$

[23]Extension to stochastic posterior beliefs is possible but not pursued. Such extension delivers similar qualitative insights with extra mathematical complications, see Cooper and Ross (1984) for useful discussion.

[24]Field experiment design is compatible with assuming away general equilibrium effects. In particular, the experiment utilizes minor idiosyncratic deviations from equilibrium price, which are unlikely to ripple in general equilibrium. Moreover, the customers are unaware of the experiment, so they are unlikely to update their belief schedule in the short run, even if the deviations were substantial and systematic.

The effect depends on $G'(p)$, which embodies the notion that "...a high price might lead some Guests to assume that listing is luxurious." The effect depends on the impact of price on market share. A price increase may lead to a more significant market share in a somewhat extreme case if $G'(p) < 0$. In other words, the marginal consumers update their beliefs about quality faster than the price increase. Such upward-sloping demand may occur for some luxurious accommodations that rely predominantly on the price to signal their quality. In that case, the negative impact of prices on reviews is guaranteed. The negative effect should be particularly pronounced in "value" review frames that put more relative weight on consumer surplus.

In another case, when $1 > G'(p) > 0$, the effect of beliefs mutes customer selection, as compared with the case of complete information. In this instance, we may observe the negative impact of price on reviews, even if the impact would have been positive with complete information.

The case $G'(p) > 1$ cannot occur at p such that $Q(p) = q = 1$. To see that examine the inequality by plugging the formula for $G'(p)$ to obtain

$$\frac{Q(p) - pQ'(p)}{Q^2(p)} > 1.$$

The inequality implies a contradiction of $Q'(p) < 0$.

The immediate observation is that asymmetric information introduces additional negative pressure of price on reviews, which may either guarantee an adverse effect or flip the direction of the effect in more marginal cases. The possible interactions of private information and the negative impact of price hikes on reviews rationalize why platforms collect and communicate the "value" frame and how this frame may be diagnostic of overpricing.

4 Data

The data was obtained from Keybee, a company that manages short-term rental properties. In comparison to long-term rentals, short-term rentals require significantly more resources to manage. Extra requirements may include replying to daily Guest messages, calendar management across platforms, scheduling turn-overs, and pricing. The pricing infrastructure of STR platforms enables elaborate dynamic pricing that reflects the current demand for accommodation. For example, the Host can set the nightly rate separately for each calendar date and change these prices as the reservation date approaches. The complexity of this decision is similar to Airline pricing.

To explain the pricing further, consider a fixed day of the year, say December 1st, 2020. Guests looking for 1-night reservations from December 1st to December 2nd may arrive at the platform any time before December 1st. Under full dynamic pricing, the price of this given reservation would vary depending on the date of the reservation query. For example, a consumer searching for December 1st on November 1st may see a different price than a consumer searching for the same date on November 2nd. Thus, pricing data is a triple panel reflecting a listing, reservation date, and price-setting date. In the remainder of the paper, we would refer to this triple as *pricing event*.

One of the services offered by Keybee is dynamic programmatic pricing. The proprietary algorithm performs daily updates of nightly rates for each unrented calendar day, that is within the next 2 years of today. The updates reflect three main factors: (i) the change in residual demand across listings, reservations dates as the reservation date approaches, (ii) decreasing opportunity cost as the reservation date approaches, and (iii) the patterns of the calendar utilization.[25] Conveniently, Keybee's pricing algorithm contains an exploration module, which perturbs the daily price of the rental by a random multiplicative shock distributed uniformly on the interval $[90\%, 110\%]$. Moreover, the perturbation is identically and independently distributed (IID) across all pricing events. In other words, nightly rates are subject to a random pricing shock that varies daily over time.

Obtaining a causal impact of a price on subsequent review is usually tricky. Prices are set as a function of cross-sectional and over-time demand shifters, which may jointly determine product reviews. For example, more luxurious vacation rentals are more expensive and obtain better reviews. Similarly, vacation rentals on the beach are more expensive in the summer and may obtain better reviews when it is warmer. This simultaneity may posi-

[25]The calendar date that has open adjacent dates has different market potential than calendar dates that do not have open adjacent dates. The lower value of monetizing gaps in the calendar occurs because the single-night (or N-night) gaps do not show in search queries for reservations longer than one (or N) night(s).

tively correlate prices and reviews, obfuscating causal associations. Conveniently, Keybee's exploration pricing provides us with the field experiment that separates these two effects.

4.1 Descriptive statistics

	Mean	SD	Min	Max	N
Total nightly fees	617.48	872.99	19.00	8,767.00	2,931
Price per night	277.51	370.23	19.00	2,922.33	2,931
Log-price per night	5.14	0.91	2.94	7.98	2,931
Number of nights	2.29	0.65	1.00	3.00	2,931
Bathrooms	1.86	1.10	1.00	8.00	2,931
Bedrooms	2.36	1.43	0.00	7.00	2,931
Review score (1-5)	4.77	0.67	0.00	5.00	2,931
Value score (1-5)	4.71	0.70	1.00	5.00	2,921
Check-in score (1-5)	4.90	0.46	1.00	5.00	2,921
Accuracy score (1-5)	4.81	0.62	1.00	5.00	2,923
Location score (1-5)	4.85	0.51	1.00	5.00	2,921
Communication score (1-5)	4.88	0.52	1.00	5.00	2,923
Cleanliness score (1-5)	4.81	0.60	1.00	5.00	2,923
Price random shock, 1st night	1.00	0.05	0.90	1.10	2,931
Price random shock, 2nd night	1.00	0.04	0.90	1.10	2,931
Price random shock, 3rd night	1.00	0.03	0.90	1.10	2,931
Low-price Arm	0.22	0.42	0.00	1.00	2,931
High-price Arm	0.16	0.37	0.00	1.00	2,931

Table 1: Descriptive statistics. The observation is a reservation. The sample consists of reservations shorter than 4 nights and receives a Guest review. The total number of observations is 2,931; however, Guests did not leave a detailed review for a small subset of reservations. For reservations shorter than n-nights, the multiplicative nightly price shock for n-th and higher nights is set to 1, which results in decreasing dispersion. High-price (low-price) Arm is defined as either: (i) a reservation for which at least two shocks are strictly greater (lower) than 1, or (ii) a 1-night reservation with a shock strictly greater (lower) than 1.

We obtained data on listing characteristics, pricing events, reservations, and reviews from Keybee. The data contains a random selection of 287 listings. For each listing, the data contains a random sub-selection of reservations from 11/29/2018 until 08/26/2021 to obfuscate the cash flow and protect the Hosts' privacy. In addition, we excluded less than 5% of reservations that did not result in a review. Finally, we removed observations for which the reporting system did not successfully record a pricing date due to downtime of the Google Big Query engine. Finally, we consider only shorter reservations of less than 4 nights because our instrument is weak for longer reservations. A drop in statistical power arises because the noise is applied to the cost of each day independently; thus, the impact of noise on the total reservation costs averages away for longer reservations. This way, we also avoid additional noise introduced by long-term discounts. As a result, we obtain 2,931 reservations.

Table 1 presents descriptive statistics about the data set. The total reservation price is a two-part tariff. The final price consists of *cleaning fee* which is paid regardless of the number of nights booked and nightly rate. The cleaning fee cannot be set as a function of the reservation dates and is fixed for a given listing over time. In most regressions, the cleaning fee would be thus subsumed into listing fixed effects. Total nightly fees per reservation average to $617, and there is considerable variation across listings. The cheapest recorded reservation is $19, while the most expensive is nearly $9,000. The average nightly rate is $277, and the average reservation length is 2.3 nights.

The data set contains a variety of listings. The average number of bedrooms is 2.36; however, we have data on studios (0 bedrooms) and 8 bedrooms listings.

The second panel of Table 1 contains average review values. The public overall review score averages 4.77. According to our sources inside Airbnb, this is somewhat larger than average review on the platform.[26]

The remaining review frames deliver higher average review ratings than the overall rating. The overall review rating is closely related to the lowest framed score, which indicates the area of deficiency.

The third panel of Table 1 contains the descriptive statistics of the pricing experiment. As mentioned above, the data contains reservations shorter than 4 nights; thus, each reservation

[26]Our data set has a larger-than-average review score than Airbnb because most of our listings are mostly professional Hosts, who typically receive better satisfaction scores, and because outsourcing management tasks to Keybee generally increases the service quality. As a result, our estimates are conservative because they are subject to ceiling effects as described by Everitt and Skrondal (2010).

is subject to at most three pricing shocks. The shock ζ_t denotes multiplicative shock to the t-th reservation date. If the reservation date is shorter than 3 nights, we set the absent shocks to 1 to balance the panel.[27] Setting missing shocks to 1 explains why the standard deviation of ζ decreases as the number of reservation nights increases.

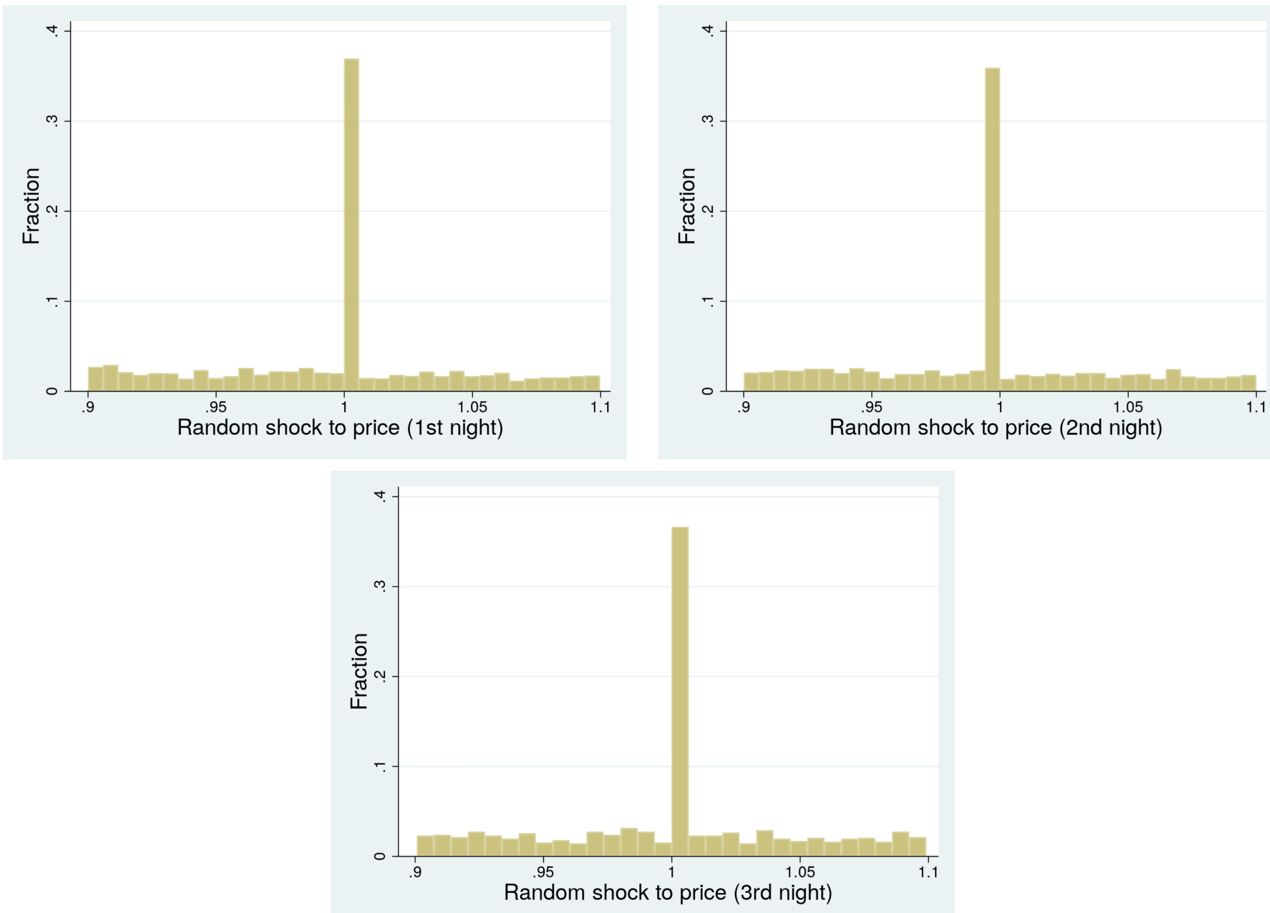


Figure 3: Distribution of the exploratory random price shocks in the population. Second-day nightly fee graph excludes 1-day reservations. Similarly, third-day nightly fee graph excludes 1-day and 2-day reservations. One-third of pricing events were excluded from the experiment at random which results in a large mass at 1.

Figure 3 contains a histogram of pricing shocks. The histogram exhibits the mass at 1, reflecting that approximately one-third of pricing events are randomly excluded from the exploratory module. We notice that the pricing shocks are uniformly distributed. To obtain model-free evidence, it is useful to consider two pricing arms. We define a low (high) pricing arm as reservations that have 2 or more shocks that are negative (positive) and single-night

[27]Pricing shocks for adjacent dates may affect the reservation demand since Guests may be choosing between stay lengths. We repeated the analysis using the actual pricing shocks for the missing dates instead of 1 and obtained the same results.

reservations with negative (positive) pricing shocks. On average, 22% of reservations are in the low pricing arm, and 16% are in the high pricing arm. The difference reflects the downward sloping demand for reservations. The selection of consumers when price varies and the corresponding selection of posted reviews is compatible with our theoretical model.

	(1)	(2)	(3)
	epsilon1	epsilon2	epsilon3
Bathrooms	-0.00192	0.000611	0.000925
	(0.00142)	(0.00135)	(0.000888)
Bedrooms	0.000570	0.000581	-0.000121
	(0.00109)	(0.00104)	(0.000683)
N	2931	2931	2931

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Randomization checks

Table 2 contains randomization checks. We regress each ζ shock on the number of bedrooms and bathrooms in the listing to obtain the coefficients. We find that none of the coefficients is statistically signification. We also performed a t-test across pricing arms and obtained similar results. In the next section, we present model-free and model-aided results.

5 Results

This section provides evidence of unilateral price changes on subsequent reservation reviews. We examine impact across several review frames with a particular focus on the *overall rating* and *value rating* frames. The former frame is the only public review frame; thus, it should be the only one that directly affects subsequent demand. The *value frame* is relevant to testing theory since it captures the effect of price hikes on reviews via consumer surplus. We start that presenting model-free evidence, which is a series of t-tests.

5.1 Model-free evidence

Figure 4 contains a comparison of histograms of reservation reviews across pricing arms. We note that high-price arm reservations obtain a smaller fraction of 5-star reviews. The

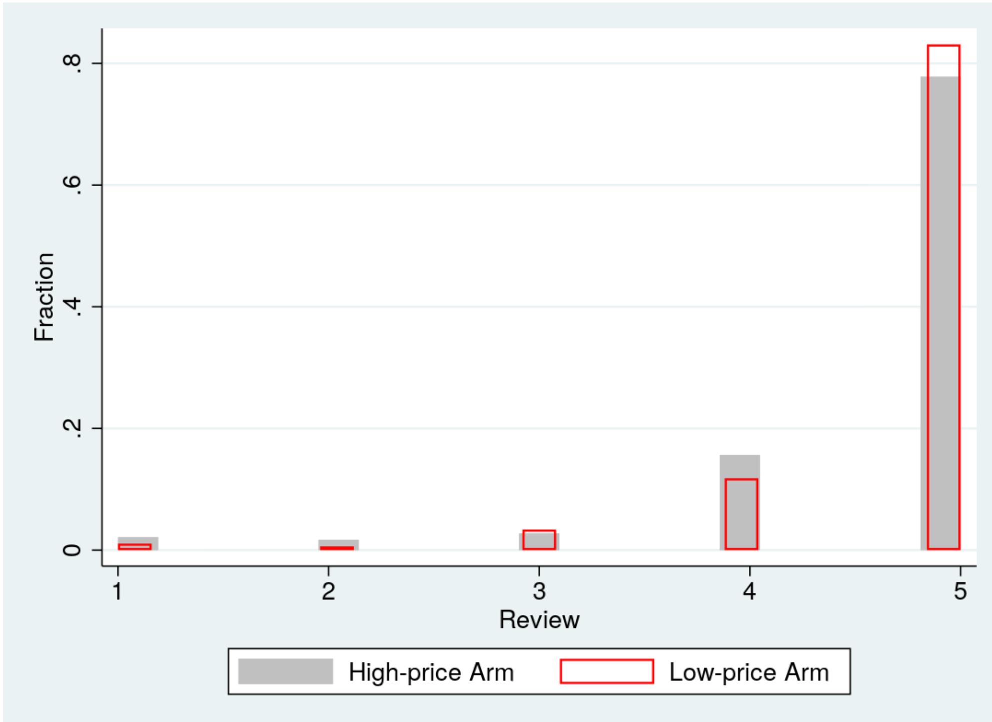


Figure 4: Histogram of reservation reviews across pricing arms.

displaced 5-star reviews are distributed across 1- to 4-star reviews, with the highest increase among 4-star reviews. We also observe a noticeable increase of very low 1- and 2-star reviews in the high-price arm. This increase is potentially the most damaging to Hosts because it effectively precludes them from obtaining a *Superhost* badge.[28] Host obtains a superhost barge if their average rating is 4.8 and higher in the prior year. Obtaining merely one 2-star review requires an offset of at least 14 subsequent 5-star reviews to maintain the badge.

Next, we conduct a t-test of review means across pricing arms. The results are presented in Table 3. The average rating drops by on average 0.9 in the high pricing arm. Further, the value rating drops by a comparable amount, suggesting lower consumer surplus is a relevant mechanism behind lower overall scoring. Interestingly, almost all other quality frames, except for *check-in* frame, reflect statistically comparable impact. These similarities suggest that guests apply a uniformly different scoring across various quality dimensions depending on

[28]There are numerous ways in which the demand is affected by the Superhost badge. Primarily, Airbnb search ranking is significantly affected by the badge. Second, the customers may filter the search results by removing non-Superhosts, which is particularly detrimental in periods of low demand. The third and perhaps most apparent factor is that the barge is displayed an icon depicting a medal next to the Host's profile photo.

	(1) Guest review score (1-5)	(2) Guest Value score (1-5)	(3) Guest Check-in score (1-5)	(4) Guest Accuracy score (1-5)	(5) Guest Location score (1-5)	(6) Guest Communication score (1-5)	(7) Guest Cleanliness score (1-5)
High-price Arm	-0.0900**	-0.0998**	-0.0000454	-0.0862**	-0.0799**	-0.142***	-0.0680**
	(0.0418)	(0.0432)	(0.0260)	(0.0402)	(0.0343)	(0.0359)	(0.0340)
N	1132	1125	1125	1127	1125	1127	1127

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Model-free evidence

the price. One consistent mechanism that rationalizes these effects is formalized in our model with asymmetric quality information. Consumers may infer the quality of, say, Guest communication or cleanliness from the posted price, which leads to a different selection of Guests in the high-pricing arm and possible ex-post regret.

A handy frame to gauge asymmetric information is the *location* frame, which aims at measuring description accuracy (as communicated to the Guests during the review process). Since listing description itself is unrelated to pricing shock, it must be that Guests update their priors about quality using both listing descriptions differently depending on the posted price. To provide a quantitative assessment of the degree of the impact, we build a regression model relating price to reviews.

5.2 Model-aided evidence

Consider the following reduced form of our theoretical review model

$$r_{it} = \delta_i + \beta X_{it} + \gamma \log p_{it} + \epsilon_{it}, \tag{1}$$

where r_{it} is the star rating, δ_i is the listing fixed effect, X_{it} are time covariates such as day of the week, month, and year of the reservation, but also characteristics of the time that the reservation was booked, such as a number of the days to the reservation, and a day of the week of the booking date. The parameter γ embodies the causal impact of price change on reviews. We opt for a log-price model since the experiment was multiplicative, and we will use a 2-stage least squares estimator.

Prices p_{it} are correlated with unobserved demand shifters. Simultaneously, demand

	(1) Guest review score (1-5)	(2) Guest review score (1-5)	(3) Guest review score (1-5)	(4) Guest review score (1-5)
Log-price	0.122***	0.116***	-0.0273	-0.0859**
	(0.0116)	(0.0120)	(0.0339)	(0.0350)
Date Controls	no	yes	yes	yes
Listing Controls	no	no	yes	no
Listing FE	no	no	no	yes
Experiment	no	no	no	no
N	2929	2929	2924	2929

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Without experiment

shifters are likely correlated with unobserved review shifters ϵ_{it}. For this reason, the baseline model suffers from endogeneity bias. Nevertheless, we start by estimating a baseline model, using a regression of star rating on log price without any control or instruments. The coefficient is presented in Column (1) of Table 4. As a result, we obtain a positive and statistically significant correlation between log price and reviews. The coefficient on the log price amounts to .122, or approximately 0.2 standard deviation increase in reviews when the log price increases by a single standard deviation.

As a first step to alleviate endogeneity bias, we add date controls generated by calendar-date specific demand shifters. As a result, we observe an economically and statistically negligible decrease in the price coefficient. Further, we add listing controls such as the number of bedrooms and zip-code fixed effects. We note a more substantial decrease in γ, suggesting that cross-sectional demand shifters generate more substantial endogeneity than calendar-day demand shifters. Finally, we add the listing fixed effect to control for time-persistent unobserved heterogeneity. Specification with listing fixed effects is the most robust model that does use experimental variation. In this model, increasing the price by one standard deviation within the listing leads to a 0.12 standard deviation decrease in star rating. The decrease in a star rating is substantial and, for example, enough to remove the superhost badge.

Next, we introduce a fully causal analysis using experimental price variation. In partic-

	(1) Guest review score (1-5)	(2) Guest review score (1-5)	(3) Guest review score (1-5)
Log-price	-1.602**	-0.851*	-0.969**
	(0.808)	(0.496)	(0.467)
Date Controls	yes	yes	yes
Listing Controls	yes	no	no
Listing FE	no	yes	yes
Experiment	yes	yes	yes
Without single-night	no	no	yes
N	2924	2929	2605

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: With experiment

ular, we consider various specifications in which we use pricing shocks ζ as instruments for log-price. The results are presented in Table 5. In addition, we consider specifications with and without various controls. We also consider removing single-night reservations, which exhibit a minor variation in reviews. Unfortunately, restricting the price variation to the pricing shocks carries a significant statistical cost. As a result, the point estimates are noisier; however, they remain significant at a 0.05 level, indicating a causal relationship between price and reviews. The point estimates indicate that increasing log price by one standard deviation leads to between 1.1 and 2.2 standard deviation decrease in reviews depending on the specification. Since standard deviation is mostly driven by cross-sectional price variation, another, possibly more informative measure, is to increase the price by a fixed proportion, say 25%. Such price increase increase leads to decrease in reviews between 12% and 23% of their standard deviation. All in all, the gap between these estimates and regressions with fixed effects suggests that time-varying demand shifters cause a significant bias when using panel data variation in prices and reviews.

We also examine the impact of log price on the framed reviews using the IV approach. We confirm that price has the largest impact on the overall score compared with framed reviews. We also find impact of the price on the *value*, *accuracy* and *guest communication* reviews. The impact of the *value* score is again of comparable magnitude to the *overall*

	(1) Guest review score (1-5)	(2) Guest Value score (1-5)	(3) Guest Check-in score (1-5)	(4) Guest Accuracy score (1-5)	(5) Guest Location score (1-5)	(6) Guest Communication score (1-5)	(7) Guest Cleanliness score (1-5)
Log-price per night	-1.606**	-1.366*	0.221	-1.857**	-0.663	-1.560**	-0.713
	(0.810)	(0.760)	(0.381)	(0.857)	(0.508)	(0.731)	(0.566)
Date Controls	yes		yes	yes	yes	yes	yes
Listing Controls	yes	yes	yes	yes	no	yes	yes
Listing FE	no	no	no	no	no	no	no
Experiment	yes	no	yes	yes	yes	yes	yes
N	2926	2916	2916	2918	2916	2918	2918

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Individual reviews

score, which provides further evidence for the consumer-surplus-based scoring rule detailed in the theoretical section. Nevertheless, we also find a significant impact on the *accuracy* score, which supports the mechanism relying on asymmetric information.

	(1) Guest review score (1-5)	(2) Guest Value score (1-5)	(3) Guest review score (1-5)	(4) Guest Value score (1-5)
Log-price per night	-0.230***	-0.143*	-0.277***	-0.218**
	(0.0835)	(0.0870)	(0.0912)	(0.0990)
Log-price× above median description length	0.178*	-0.00196		
	(0.0912)	(0.107)		
Log-price× above median picture count			0.241**	0.119
			(0.0993)	(0.116)
Date Controls	yes	yes	yes	yes
Listing Controls	no	no	no	no
Listing FE	yes	yes	yes	yes
Experiment	no	no	no	no
N	2082	2072	2082	2072

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Individual reviews

Our theoretical model demonstrates that asymmetric information is not necessary to obtain the negative impact of transaction price on the subsequent product reviews. That being said, the information asymmetries may, in some cases, amplify this effect. To test this hypothesis, we estimate heterogeneous treatment effects based on the amount of revealed information. In particular, for the subset of listings, we obtained information about the length of the listing description and picture count. Therefore, we presume that longer descriptions and more pictures result in a relatively minor degree of asymmetric information. Unfortunately, we could not obtain this data for all listings because a subset of Hosts stopped hosting during the data collection process. For this reason, to conduct this exercise, we have to revert to the OLS model with listing fixed effects due to the insufficient power of the IV estimator. For these reasons, the results should be regarded as conservative.

The results of the heterogeneous treatment effect regression are presented in Table 7. We observe that the effect of pricing on overall review is significant for listings with less-than-median description length and picture count. However, we observe significantly smaller effects for listings with lengthy descriptions and many photos. The muted effect of the transaction price of reviews for the listings with arguably more accurately unobserved quality suggests that the theoretical model of review deliberation with asymmetric information may be more realistic.

Consistent with our theoretical model, an alternative explanation for the heterogeneous treatment effects is that review length and pictures serve as a proxy for listing quality. As noted in the theoretical section, the impact of pricing on customer feedback may be less extensive for high-quality listings because the Guests in such listings may be less price-elastic and focus more on pure quality instead of consumer surplus. An extreme example may be customer satisfaction from staying in a multi-million-dollar mansion, which attracts only extremely wealthy Guests, who may not even be aware of the transaction price. In this case, however, we should see a more muted impact of transaction price in the *value* frame, since under the heterogeneous quality hypothesis value frame becomes less relevant for wealthy consumers. Interestingly, we do not find heterogeneous treatment effect in *value* frame.

In the next subsection, we conduct a series of placebo tests using the time-series feature of our data.

5.3 Placebo tests

The placebo tests are designed to rule out any spurious correlation due to an error in the pricing and randomization algorithms or selection of observations due to downtimes in the data collection scripts. To conduct the test, we pulled the pricing arm of the previous

	(1) Guest review score (1-5)	(2) Guest Value score (1-5)	(3) Guest Check-in score (1-5)	(4) Guest Accuracy score (1-5)	(5) Guest Location score (1-5)	(6) Guest Communication score (1-5)	(7) Guest Cleanliness score (1-5)
Previous reservation High-price Arm	-0.0531	-0.00108	-0.00439	0.00350	0.0259	0.00544	-0.0266
	(0.0428)	(0.0426)	(0.0306)	(0.0389)	(0.0312)	(0.0300)	(0.0374)
N	1387	1379	1379	1381	1379	1381	1381

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: With experiment

observation for a given listing and used it as a covariate in the review t-tests instead of the shock to the current reservation. Since the current Guest does not observe the previous reservation price, it should have no relationship with the current review. Indeed, we find no significant relationship between the previous pricing arm and the current review score across all review frames.

In the next section we conclude our analysis with parsimonious analysis of the supply-side impact of reviews on prices.

5.4 Supply

	(1) Log-price per night	(2) Log-price per night	(3) Log-price per night	(4) Log-price per night	(5) Log-price per night
prev_score	0.254***	0.214***	0.0264***	0.00648	1.458**
	(0.0136)	(0.0133)	(0.00524)	(0.00501)	(0.575)
Date Controls	no	yes	yes	yes	yes
Listing Controls	no	no	yes	no	no
Listing FE	no	no	no	yes	no
Experiment	no	no	no	no	yes
N	6187	6187	6182	6187	6187

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Price cycles

In this subsection, we examine the impact of reviews on supply. There are two goals of this analysis. First, we want to estimate the causal impact of reviews on the subsequent transaction price. Second, we aim to identify data patterns consistent with price cycles, similar to predatory-dominant pricing patterns as in markets with learning by doing, see Cabral and Riordan (1997).

Airbnb uses various ways to communicate listing reviews to potential Guests. Next to the listing's name, the STR platform displays the average review for that listing. In addition, after clicking on the listing name, Airbnb provides a chronological reviews list, displaying star ratings and review content. The chronological display is aimed the highlighting the most recent reviews since those may be more reflective of the contemporary listing quality.

Relating past review to the price faces an obvious simultaneity issue since higher quality listings receive higher reviews and are more expensive. In addition, simple review regression on transaction prices is subject to reverse causality because of the direct impact of pricing on reviews demonstrated in the previous section. Our analysis would be robust to both of these issues.

Before we introduce experimental variation, we estimate a model that utilizes a within-listing review and transaction price variation. This model would address cross-sectional simultaneity caused by persistent unobserved heterogeneity in listing quality. We estimate estimate the following supply model

$$\log p_{it} = \delta_i + \beta X_{it} + \gamma r_{it-1} + \epsilon_{it}, \tag{2}$$

where i indexes listings and t indexes reservations within the listing. The variable r_{it-1} is the previous rating displayed on the top of the listing page. We do not use the average star rating because it is quite sticky over time since most of our listings have many reviews.

The supply equation (5.4) is the reverse of the review deliberation equation (5.2). In particular, previously, we used reviews as an independent variable and price as a dependent variable. In this analysis, we study the impact of reviews on the price. Another crucial difference is that the supply model uses lagged reviews on the right-hand-side.[29]

The results are presented in Table 9. The first row presents the regression without any controls. We observe a strong correlation between past reviews and future transaction prices. Row (2) addresses simultaneity due to seasonality. In particular, certain times of the year, such as Summer, may be subject demand shocks driving higher prices and directly impact-

[29]Before we discuss the results, we disclose that this analysis was conducted using a different selection of reservations as compared to the results in the previous section. The previous analysis selected reservations shorter than 4 nights since the pricing instruments are weak for longer reservations. We use lagged ratings for reservations shorter than 4 nights anticipating similar issues.

ing review scores. We observe that the coefficient decreases by a negligible amount. As mentioned before, the high value of the coefficient may be due to cross-sectional unobserved heterogeneity. Models whose coefficients are in Rows (3) and (4) address the unobserved heterogeneity issue by controlling for listings characteristics and listing fixed effects, respectively. We see that the effect persists and is statistically significant. Nevertheless, these regressions may still be problematic because of using lagged dependent variables with fixed effects. Also, we would still be subject to simultaneity if ϵ_{it}s are serially correlated. According to our anecdotal knowledge, the latter is probable since listings experience transient quality shocks. For example, a broken heater may be an issue for several subsequent reservations until it is finally fixed. We use the fact that our experimental pricing variation is IID across pricing events to address this issue. Using equation (5.2), we established that the current transaction price affects reviews. Thus, lagged pricing shocks, ζ_{it-1} are correlated with lagged reviews, r_{it-1}. At the same time, pricing shocks are IID; thus, they are unrelated to contemporary unobserved determinants of price, ϵ_{it}. In particular, lagged pricing shocks are orthogonal to transient shocks to unobserved quality and contemporary pricing shocks, ζ_{it}. As a result, the experiment design presents a unique opportunity to provide causal estimates of γ robust to both simultaneity and reverse causality.

The robust estimate of γ is presented in Row (5). The coefficient is significant at the 5% level, albeit its point estimate is quite noisy. We use lower 95% confidence interval value to obtain a conservative estimate of the impact of prices on reviews. The lower confidence bound is equal to 0.33, per 1 star, or 0.22 per one standard deviation of review score. This coefficient implies 25% increase in transaction price if the last review improves by 1 standard deviation. The change is due to the earlier booking time at higher dynamic price. As mentioned earlier, since the listings eventually reach over 90% filled capacity, the change in transaction price reflects change in revenue to high degree of accuracy. Large impact of last review on the performance suggests that the incentive to invest in great reviews by lowering the price is substantial.

If the current transaction price has a causal impact on reviews, firms have incentives to lower their price after a bad review to invest in the future good reviews. But, conversely, they also have incentives to increase the price and cash out if the reviews are excellent. The estimates in Table 9 are consistent with such price cycles. The coefficients are also consistent with optimal static pricing that responds demand shifters, including past reviews. Unfortunately, reduced form modeling is insufficient to decompose static pricing from dynamic pricing incentives. For example, the Keybee pricing algorithm is based on forward-looking machine learning routines that maximize Hosts' revenue.[30] The algorithm automatically estimates and responds to shocks to demand. Unfortunately, similarly to most machine learning, the algorithm is not economically explainable and may require a structural model for further understanding.[31]

There are several consequences on the incentive to price low in order to buy honest reviews and increase future demand. Primarily, some increase in future demand would be a result of business stealing. Prior research has demonstrated some degree of supply-side substitution between regular and short term rental market;[32] thus, one can expect exit from short-term rental market in case of low revenues, which could occur due to lower ratings. The possibility of exit gives extra incentives of the firm to drop their prices to preempt entry or drive exit. These incentives should be more pronounced in thinner markets and could lead to predatory pricing and lower competition in the longer run, similarly to markets with learning-by-doing. On the other hand, price-review feedback loop gives consumers an ability to recourse, if the margins increase substantially. The potential recourse creates a downward pricing pressure, which is present even in completely monopolized markets. Thus, consumer welfare implications of the ability buying honest reviews are unclear.

[30]Casual Hosts that do not apply sophisticated dynamic pricing may be less prone to price cycles; that being said, setting even the most basic dynamic pricing, such as "early-bid" and "last-minute" discounts, which are easily accessible in Airbnb Hosting Dashboard, could lead to price cycles.

[31]A possible test for investment cycles could be studying heterogeneous treatment effects of past reviews on future prices depending on the degree of asymmetric information. Theory suggests that investment price cycles should be muted for listings with less private information. Unfortunately, our instruments do not directly manipulate reviews and rely on indirect impact via past transaction prices; thus, they do not deliver statistically robust estimates of heterogeneous treatment effects. We speculate that augmenting the regressions with a structural model could deliver necessary statistical power.

[32]See Barron, Kung, and Proserpio (2018).

6 Conclusion

The paper uses a field experiment to establish the causal impact of transaction price on subsequent product reviews. Despite the central position of reviews in modern markets, surprisingly little is known about the customer deliberation process that results in a rating. In particular, little is known if the ratings reflect consumer surplus or raw quality. We also know little about the interaction of asymmetric quality information and subsequent review.

The paper finds that increasing the price has causal negative impact on subsequent review. Conversely, current high reviews have causal positive impact on future transaction prices. These results provide evidence that consumer surplus and asymmetric information about the quality are essential considerations in the consumer post-purchase deliberation process. The importance of consumer surplus has significant consequences. Primarily, the reviews must be evaluated in the context of past prices; otherwise, they would not provide an apples-to-apples comparison of quality across products and within-product over time.

The identified effects have consequences for market efficiency. On the one hand, because past reviews influence future company profits, the firms face downward pricing pressure. The pressure occurs because consumers obtain a way to punish the firm if the transaction price seems excessive compared to offered value. The effect is present even for a monopolist, demonstrating that the rating systems can serve as an essential competitive force even in the market places with concentrated ownership. On the other hand, because the upcoming competitor has to invest in good reviews, it is conceivable that good reviews of the incumbent would deliver, at least temporary, market power that facilitates market dominance.

References

Akerlof, G. A. (1978): "The market for "lemons": Quality uncertainty and the market mechanism," in *Uncertainty in economics*, pp. 235–251. Elsevier.

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

Arellano, M., and S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The review of economic studies*, 58(2), 277–297.

Barron, K., E. Kung, and D. Proserpio (2018): "The sharing economy and housing affordability: Evidence from Airbnb," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, pp. 5–5.

Bondi, T., and R. Stevens (2019): "The Good, The Bad and The Picky: Consumer Heterogeneity and the Reversal of Movie Ratings," .

Borkovsky, R. N., A. Goldfarb, A. M. Haviv, and S. Moorthy (2017): "Measuring and understanding brand value in a dynamic model of brand management," *Marketing Science*, 36(4), 471–499.

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

Cabral, L. M., and M. H. Riordan (1994): "The learning curve, market dominance, and predatory pricing," *Econometrica: Journal of the Econometric Society*, pp. 1115–1140.

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

Chevalier, J. A., and D. Mayzlin (2006): "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 43(3), 345–354.

Cooper, R., and T. W. Ross (1984): "Prices, product qualities and asymmetric information: The competitive case," *The Review of Economic Studies*, 51(2), 197–207.

Dubé, J.-P., G. J. Hitsch, and P. Manchanda (2005): "An empirical model of advertising dynamics," *Quantitative marketing and economics*, 3(2), 107–144.

Everitt, B. S., and A. Skrondal (2010): "The Cambridge dictionary of statistics," .

Fradkin, A., E. Grewal, and D. Holtz (2021): "Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb," *Marketing Science*, 40(6), 1013–1029.

Hu, N., P. A. Pavlou, and J. J. Zhang (2017): "On self-selection biases in online product reviews.," *MIS Q.*, 41(2), 449–471.

Laouénan, M., and R. Rathelot (2022): "Can information reduce ethnic discrimination? evidence from airbnb," *American Economic Journal: Applied Economics*, 14(1), 107–32.

Lewis, G. (2011): "Asymmetric information, adverse selection and online disclosure: The case of eBay motors," *American Economic Review*, 101(4), 1535–46.

Li, X., and L. M. Hitt (2008): "Self-selection and information role of online product reviews," *Information Systems Research*, 19(4), 456–474.

Luca, M., and O. Reshef (2021): "The effect of price on firm reputation," *Management Science*, 67(7), 4408–4419.

Stenzel, A., C. Wolf, and P. Schmidt (2020): "Pricing for the stars: Dynamic pricing in the presence of rating systems," in *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 273–274.

Wolinsky, A. (1983): "Prices as signals of product quality," *The review of economic studies*, 50(4), 647–658.

Zhong, Z. (2021): "Chasing diamonds and crowns: Consumer limited attention and seller response," *Management Science*.