The Supply and Demand Effects of Review Platforms[†]

Gregory Lewis
Microsoft Research

Georgios Zervas
Questrom School of Business
Boston University

October 14, 2019

Abstract

Review platforms such as Yelp and TripAdvisor aggregate crowd-sourced information about users' experiences with products and services. We analyze the impact of review platforms on the hotel industry using a panel of hotel prices, sales and reviews from five US states over a 10-year period from 2005-2014. We show that a hotel's demand is positively correlated with their average rating across TripAdvisor, Expedia and Hotels.com. Such correlations have grown over our sample period from a statistical zero in 2005 to a substantial level in 2014: a hotel rated one star higher on all the platforms on average has 27.8% higher demand. By contrast, an increase in average ratings has a negligible effect on hotels' pricing schedules (i.e. supply). A natural experiment in our data that caused abrupt changes in the ratings of some hotels but not others suggests that these associations are causal. Building on this causal interpretation, we estimate heterogeneous treatment effects: independent hotels stand to gain more from online reputation than chains, as do hotels that practice revenue management.

[†]We thank Michael Butler and Michael Kurish for excellent research assistance. This paper is part of a larger project, and we are grateful for comments on this project from James Dana, Ginger Jin, and Ryan McDevitt, as well as seminar participants at Airbnb, Amazon, Chicago, Columbia, DOJ, Duke, EC, Marketing Science, Maryland, Michigan, MIT, NBER, NYU, Princeton, QME, Stanford, Toronto, USC, UT Austin, and Wharton. We also thank STR, and in particular Duane Vinson, for providing data.

1 Introduction

Consumers are increasingly turning to the internet to learn about which products to buy. One of the main information sources they consult is peer reviews, which are now ubiquitous on online platforms such as Amazon, Google, TripAdvisor and Yelp. A large body of data has accumulated: TripAdvisor — the world's largest travel review platform — contains 320 million reviews, while Yelp — which aggregates reviews for local businesses — receives approximately 150 million visitors per month.[1] Using regression discontinuity designs (Hahn et al., 2001; Lee and Lemieux, 2010), previous work has shown that online reviews have a positive effect on business revenue (Luca, 2011; Anderson and Magruder, 2012). This leaves open a number of intriguing questions: *How do online reviews impact hotel supply and demand? Have these effects changed over time? How does the importance of reviews vary with organizational type and management practices?* These are all relevant factors for managers as they need to understand when to invest in reputation, and how best to capitalize on it.

To answer these questions, we put together a novel and extensive dataset. It merges a decade-long panel data from Smith Travel Research (STR), the largest data-provider in the hotel industry, on hotels in the western United States with data that we scraped on the reviews from three major hotel platforms: Expedia, Hotels.com and TripAdvisor. Because average prices are reported to STR, we can disaggregate revenue into prices and quantities. This allows us to look at how supply and demand have responded to reviews, and how these effects vary across time and organization.

We uncover a number of intriguing patterns. The number of online reviews is increasing and actually accelerating, due to an increased rate at which consumers leave reviews. The fidelity of average reviews, as measured by the correlation in average reviews across platforms, is also increasing. Both of these suggest that online reviews are increasingly informative.

Correspondingly, we show that hotel demand has become increasingly responsive over time to online ratings. Whereas online ratings had a statistically zero effect on demand in 2005, we estimate that by 2014, the last year of our data, a 1 unit increase in average ratings across all the platforms leads to a 27.8% increase in demand. Both occupancy and prices are higher, all else equal, at higher-rated hotels.

One possible explanation for higher prices is active management by hotels: hotels whose

[1] See `http://www.tripadvisor.com/PressCenter-c4-Fact_Sheet.html` and `http://www.yelp.com/factsheet`.

ratings increase raise their prices. We find little evidence of this: the relationship between prices and occupancy (i.e. the individual hotel supply curve) does not seem to vary much with changes in ratings. Instead, we find that the explanation is dynamic pricing, i.e. setting a price schedule that is increasing in the number of rooms already sold. This implies that hotels whose ratings increase charge higher prices for their marginal rooms and hence have higher average prices.

We also cut the data in a number of ways to enrich this story. We look at differences in organizational form, showing that independent hotels benefit more from higher ratings than chain hotels, consistent with online reputation being a substitute for brand familiarity. The same is true of higher-end hotels, where the estimated effects of online reputation are much bigger. This makes sense, as travelers to these hotels typically put a high premium on quality. Finally, we show that there is complementarity between pricing strategies and reputation: hotels that we estimate to be in the top half of hotels in terms of how aggressively they adjust prices in response to changes in occupancy earn on average 12% more from a 1 unit increase in ratings, as compared to 5% for those in the bottom half.

To reach these conclusions, we must overcome two different estimation challenges. The first is the simultaneous determination of price and quantity. We address this problem using tools from the industrial organization literature, identifying the slope of the demand curve through a combination of first order conditions on the supply side and demand shifters (in this case, temperature).

The second is determining the causal impact of ratings, as distinct from any omitted factor that is correlated with price, quantity, and ratings. Our main strategy is simply to add many flexible controls to account for these potential omissions, combining hotel and market-time-period fixed effects. The identifying variation then becomes the within-variation in hotel ratings that is not accounted for by market-time specific trends.

In addition to this, we analyze a natural experiment, the merger of the review databases of Expedia and Hotels.com. This merger changed hotel ratings overnight, because hotels with distinct scores on the different platforms now earned a single score based on all of the reviews. Due to the rounding of ratings to 0.1 increments, some hotels changed their score on one platform and stayed the same on the other, providing a clean change in their ratings that we can be sure is unrelated to any omitted factors.

1.1 Related work

The impact of reviews and ratings on firms' sales has been a popular subject in the marketing and economics literatures. Closely related to our work are a number of studies that estimate the impact of reviews and ratings on sales for various products and services including restaurants, books, and internet auctions (Resnick et al., 2006; Cabral and Hortascu, 2010; Luca, 2011; Anderson and Magruder, 2012). These and several other papers have focused on estimating the treatment effect of a firm's online reputation on its financial performance, consistently arriving at the conclusion that online reputation is a significant driver of sales. Our work differs from these papers in two ways: first, we take a market-level view rather than a firm-centric one, accounting for how changes in information disseminated through review platforms affect competition between firms within a market; second, because our data spans a decade — over which review platforms grew from insignificance to many consumers' primary source of information — we can examine the evolution of these treatment effects over time.

Along similar lines a number of papers measure the effect of quality disclosure on consumer choice. Jin and Leslie (2003) find that consumers respond to the disclosure of restaurant health ratings, and that restaurants respond by becoming cleaner; jointly these effects leads to a decrease in food-borne illness suggesting that health ratings had a positive impact on consumer welfare. Elfenbein et al. (2015) study eBay sellers and find that the extent to which a "top rated seller" helps attract more customers depends on how many other sellers who sell similar products also have the badge. Bai (2015) experimentally demonstrates that quality disclosure leads to increased prices and profits in market with high information asymmetries.

Closer to our work, a smaller literature has examined the impact of consumer reviews on firm behavior, and in particular firms' investments in their reputation. One strand of this literature examines the interaction between consumer reviews and firms' advertising decisions: Chen and Xie (2005, 2008) develop a model of the interactions between advertising strategy and consumer reviews, and Hollenbeck et al. (Forthcoming) empirically demonstrate that hotels advertise less when their TripAdvisor ratings go up. Along similar lines, Hollenbeck (2018) examines substitution between consumer reviews and firms' "offline" reputation, finding that with the rising prominence of review platforms, the value of hotel brand affiliation has decreased. Another strand of this literature has studied firms' incentives to commit review fraud — either to boost their own reputation, or, to damage the reputation of their competitors — finding that incentives to commit review fraud are higher in markets with a lot of competition, and when firms have a weak reputation (Mayzlin et al., 2014; Luca and Zervas, 2016). Finally, Proserpio and Zervas (2017), Chevalier et al. (2018), and Wang and Chaudhry (2018) show that firms frequently engage with consumers on review platforms by responding to their reviews. Interestingly, even though firms that respond to consumers presumably do so because they believe it has benefits, these papers arrive at different conclusions regarding the implications of directly engaging with consumers by responding. All of the above papers show that, in various ways, review platforms have not only changed how consumers make decisions, but also how firms behave in the marketplace.

Finally, our work informs the literature on measuring consumer surplus from the digital economy (Brown et al., 2002; Brynjolfsson et al., 2003; Goolsbee et al., 2006; Brynjolfsson and Oh, 2012; Pantea and Martens, 2014). A difficulty in calculating the consumer surplus of review platforms is that we need to estimate changes in both consumer and firm behavior as a function of platform entry. While we do not estimate consumer surplus in this paper, our work suggests that consumers many not be uniformly better off in the presence of review platforms: while some consumers gain from review platforms by making better choices, high quality firms respond by increasing prices offsetting some these gains.

2 Data and Setting

2.1 Data

Our primary source of data for this study is a decade long monthly panel of hotel financial performance, which we obtained from Smith Travel Research (STR). The STR panel contains 4,477 hotels located in Arizona, California, Nevada, Oregon, and Washington. Approximately 45% of all hotels that operated in these five states during our observation period reported financial performance data to STR, and are thus included in our panel. The hotels in our panel are much more likely to be affiliated with a chain than to be independent: 80% are chains or franchises, and 20% are independents. For each hotel-month, we observe the number of room-nights available, the number of room-nights sold, and the total room revenue generated. Using these three variables, we also calculate average room prices and occupancy rates over time. In addition to these time-varying covariates, our data contains a rich attribute set covering both STR and non-STR hotels: hotel location at the ZIP code level, opening and closing dates (if any), price category (from Budget to Luxury), organizational form (chain, franchise, or independent), capacity, and the square footage of any meeting and business facilities. Our data masks the identities of individual hotels.

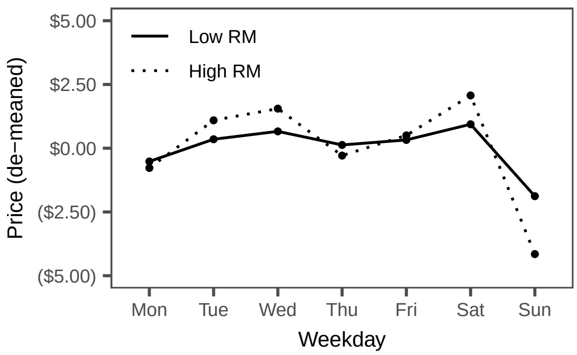
For 3,343 of the above hotels, we also have access to daily price and occupancy data, which we use to construct our revenue management proxy as described below.

We augment the hotel financial performance data with a panel of consumer reviews from three major review platforms: TripAdvisor, Expedia, and Hotels.com. In our data, we observe the entire history of 1-, 2-, 3-, 4-, and 5-star ratings for each hotel on each review platform. Our data does not contain the text of individual reviews to maintain hotel anonymity. In total, our reviews dataset contains 807,140 Expedia ratings, 1,410,488 Hotels.com ratings, and 1,544,883 TripAdvisor ratings. We aggregate ratings across platforms at the hotel-month level to match the financial performance data, defining the average rating $r_{j,t}$ of a hotel j in a specific month t as the sum of all individual ratings in reviews received by j up until time t across all three platforms, divided by the total number of reviews across the platforms.[2]

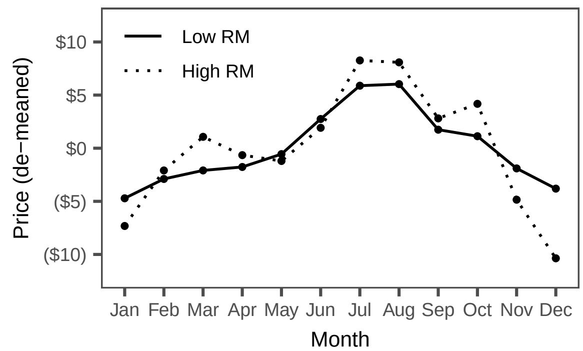
Revenue management Revenue management (RM), the practice of continuously optimizing pricing strategy as a function of forecasted demand, is prevalent in the hotel industry. However, revenue managers employ a wide array of revenue management techniques that vary in their sophistication (see Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003) for overviews of the literature on revenue management and dynamic pricing). One of our main hypotheses in this paper is that hotels that are "better" at revenue management will benefit more from increases in their ratings, by being able to adjust prices to reflect increased demand. Since we do not observe revenue management sophistication we have to rely on a proxy.

We construct a RM proxy by estimating the responsiveness of individual hotel prices to changes in occupancy, i.e. dynamic pricing. To do so, we estimate individual supply curves by regressing price on occupancy, allowing for a hotel-specific coefficient on occupancy, and adding fixed effects for day-of-week (e.g. Monday), month-day (e.g. Jan 1), year (e.g. 2013) and hotel. Our intent in adding these controls is to account for all predictable sources of demand variation that may lead hotels to systematically shift price schedules (i.e. to control for supply shifters), leaving only the "good" demand variation to identify the slope of the supply curve. The estimation procedure we use is to first condition on all the fixed effects and obtain price and occupancy residuals, and then regress the residuals against each other (i.e. a Frisch-Waugh-Lovell approach).

[2] In the appendix, we show that all our main results are robust to instead first defining platform-specific weights based on the platform share of reviews across all hotels up to time t, and then averaging average reviews on each platform according to these platform specific weights.



(a) Daily price variation.



(b) Monthly price variation.

Figure 1: Price variation in the period 2010–2014 by revenue management (RM) status estimated using 2005–2009 data.

We estimate the supply slopes on the first 5 years of data, and then use the estimated proxies in regressions on the held-out next 5 years of data, to avoid picking up any results mechanically purely as a result of the way the proxy is constructed. We omit hotels that have less than a year of data in this estimation.

We test whether our RM proxy works as intended by checking whether it predicts variation in pricing across days (i.e. day of week pricing) and months (i.e. seasonal pricing). Notice that both of these forms of price variation could arise from either a fixed time-varying schedule of prices (e.g. high season and low season prices) or from revenue management (prices are higher in the high season because the rooms in the low-priced "bucket" sell out.) Notice also that we explicitly controlled for day-of-week and month-day variation in constructing the RM proxy, so any correlation we find between the proxy and this price variation is not mechanical (also we use the held-out sample for this exercise).

Figure 1a shows the average prices by day-of-week, contrasting the prices from the low revenue management group (all hotels below the median on the RM proxy) and the high one (all hotels above the median on the RM proxy). There is more variation in the high RM group, although both show some evidence of day-of-week pricing. Figure 1b replicates this by month-of-year, and again we see more price variation in the high RM group.

Summary Statistics We summarize the resulting data in Table 1, in which an observation is a hotel. The average hotel charges $103.57 for a room, is 63% full, has a rating of 3.76 and gets 7.49 reviews per month. The mean estimated slope coefficient on occupancy is 25.37, implying that on average hotels increase prices by roughly $0.25 for each percentage point increase in occupancy rates. Later we will show that in all our main specifications, which

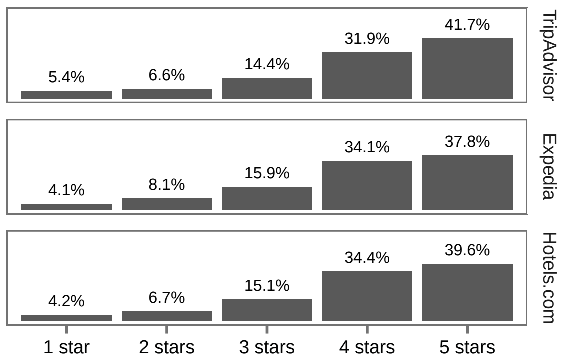


Figure 2: Distribution of ratings by review platform.

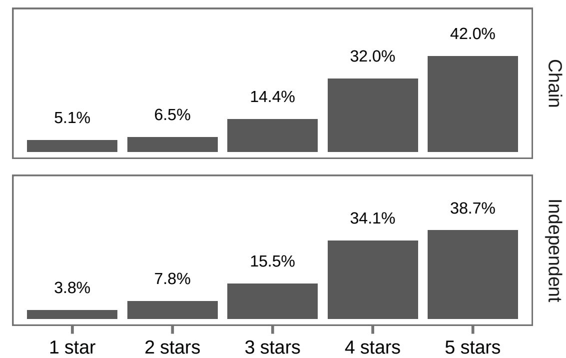


Figure 3: Distribution of ratings by hotel type.

use an instrumental variables strategy to deal with simultaneity, the estimated mean slope is substantially higher — around 70. We infer that the proxy may be downward biased for the true individual supply slopes, but since we're mostly interested in the rank-order — which hotels are most sophisticated in their revenue management — this bias is acceptable.[3]

2.2 Setting

In this section, we provide some descriptive evidence on how consumers review hotels and how those reviews influence demand. Figure 2 shows the distribution of ratings on each review platform. In our data, negative reviews are rare – only 12% of ratings are below three stars. These distributions differ from some prior measurements, which found that extreme reviews are more prevalent than moderate ones, resulting in a J-shaped distribution (Hu et al., 2009). Reviewing patterns appear to be similar across the three platforms. Over time the correlation in the average reviews across platforms has also grown, so that the platforms largely "agree" about hotel quality. For example, between 2010 and 2014, the correlation between TripAdvisor average ratings Expedia average ratings increased from 0.31 to 0.50, while the correlation between TripAdvisor and Hotels.com increased from 0.25 to 0.38. [4]

As shown in Figure 3, the distribution of ratings is also similar across chain/franchise hotels and independents, despite the fact that independent hotels tend to be of average higher price ($122 for independents versus $80 for chain or franchise hotels).

[3]We have experimented with instrumental variables strategies here too, but because we are proceeding hotel-by-hotel there is high variance in the estimates, and we prefer the biased but lower variance OLS approach.

[4]Some of this growth in correlation is mechanical, since Expedia and Hotels.com merged their review databases in June 2013. See our discussion of the natural experiment in section 4 below.

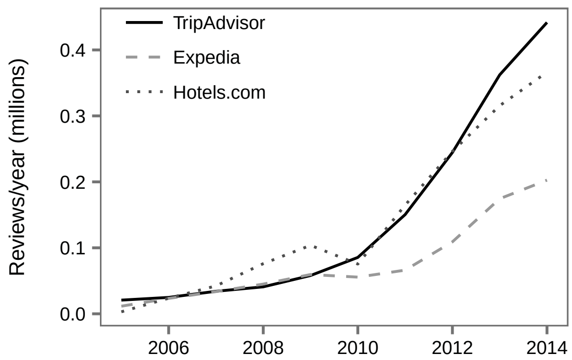


Figure 4: Number of reviews submitted by year.

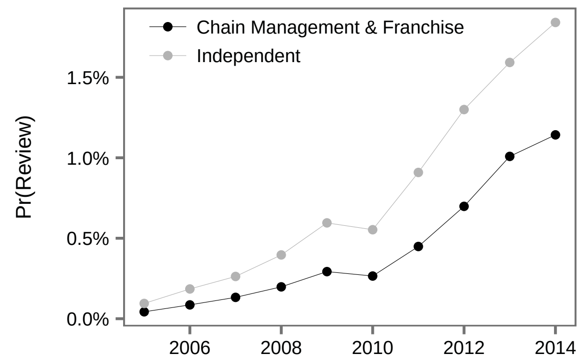


Figure 5: Probability of leaving a review on TripAdvisor, Expedia, or Hotels.com by year.

Review platforms have been growing in popularity. Figure 4 shows the number of reviews submitted by year on each platform. Notice that the growth rate has increased in recent years, particularly on TripAdvisor. The prior literature has offered little evidence on the frequency with which consumers decide to leave a review. Figure 5 plots the annual probability of leaving a review on any of the three review platform in our data, which we define as total number of reviews over total number of reservations. Because we do not observe the latter quantity, we approximate it by dividing total room-nights by the average length of stay nights, which we assume to be two nights based on industry statistics.[5] We find that, while the rate at which travelers leave reviews has been growing, by 2014 about 2% of stays at independent hotels, and 1% of stays at chain hotels resulted in a review being left on Expedia, TripAdvisor, or Hotels.com. These figures are in sharp contrast to the ones seen on platforms like Airbnb, where 67% of guests leave a review on average (Fradkin et al., 2014). In summary, the quality of information has increased over time, as measured by the number of reviews, the number of hotels covered, and increasing convergence in average ratings — and given the increasing participation rates, this trend looks set to continue.

3 Empirical Strategy

Our goal is to learn how hotel reviews affected supply and demand in the hotel market, and how this relationship has varied over time and with organizational form. We face two

[5]According to the American Hotel & Lodging Association, among leisure travelers staying at a hotel "50% spend one night, 26% spend two nights, and 24% spend three or more nights". Business trips are typically of shorter duration. These statistics suggest an average length of stay of approximately two nights. See `https://www.ahla.com/content.aspx?id=36332`.

challenges in doing so. First, as always, price and quantity are simultaneously determined. We develop an empirical strategy that resolves this simultaneity problem below in Section 3.1.

Second, hotel ratings may be endogenous: that is, correlated with unobserved information sources (e.g. marketing efforts) that drive consumer demand. In this section of the paper we address this by adding a large set of hotel and market-time-specific fixed effects to control for persistent differences across hotels, and time-varying market-level demand shocks. This is our main specification, as it leverages all of our data, giving us sufficient power to detect the heterogeneous treatment effects that are of interest.

Still, this approach will still lead to biased estimates in the case where there are within-market time-varying sources of individual hotel demand that are correlated with ratings. For example, if individual hotels increase their advertising precisely when their own reviews improve we may incorrectly attribute increased sales to the reviews rather than their advertising campaigns (though we are robust to aggregate seasonal advertising patterns, since we have market-time fixed effects). To test for this bias in our results, we take advantage of a natural experiment that occurs during our observation period: a merger between Hotels.com and Expedia that creates plausibly exogenous variation in ratings in Section 4. Previewing that section, we find no statistically significant evidence of bias, which gives us confidence in the results of our main specification.

3.1 Deriving the estimating equations

We begin by specifying the demand for a hotel:

$$o_{j,t} = \alpha p_{j,t} + \beta r_{j,t} + \gamma_j + \mu_{m,t} + \xi_{j,t} \tag{1}$$

where j indexes hotels (e.g. San Francisco Mariott), m indexes markets (e.g. San Francisco) and t indexes time periods (e.g. June 2012). $o_{j,t}$ is the occupancy of hotel j at time t, which we choose as the dependent variable to normalize across hotels of different capacities (occupancy is rooms sold $q_{j,t}$ divided by capacity κ_j). $p_{j,t}$ is the price, $r_{j,t}$ is the average rating as defined earlier in Section 2, γ_j is a hotel fixed effect to control for fixed hotel unobservables, $\mu_{m,t}$ is a market-time-period fixed effect to control for seasonal demand by market, and $\xi_{j,t}$ is an error term.[6]

[6]We also include an indicator for a hotel having no reviews and set the average review to zero for hotels which have no reviews, allowing for a potentially non-linear effect of having zero reviews. For clarity of presentation, it is omitted in what follows.

On the supply side, we begin with the first order condition (FOC) for the hotel's profit maximization problem (assuming a constant marginal cost, see below). We can write the FOC directly in terms of occupancy, since capacity is just a fixed multiplier of the profit function, and therefore cancels out:

$$\frac{\partial o_{j,t}}{\partial p_{j,t}}(p_{j,t} - c_{j,t}) + o_{j,t} = 0 \tag{2}$$

where $c_{j,t}$ is the marginal cost of a hotel room. Let us specify the following form for marginal cost:

$$c_{j,t} = \beta_s r_{j,t} + \delta_j + \mu^c_{m,y} + u_{j,t} \tag{3}$$

a linear term in ratings, a hotel specific effect, a market-year specific effect to account for inflation (y indexes year) plus an error term. Since we don't believe marginal costs vary with ratings, we expect to estimate $\beta_s = 0$. Noticing from (1) that $\frac{\partial o_{j,t}}{\partial p_{j,t}} = \alpha$, we re-arrange to get a supply equation:

$$p_{j,t} = -\frac{1}{\alpha} o_{j,t} + c_{j,t} = \beta_o o_{j,t} + \beta_s r_{j,t} + \delta_j + \mu^c_{m,y} + u_{j,t} \tag{4}$$

We estimate the supply equation (4) by two-stage least squares (2SLS) instrumenting for the occupancy $o_{j,t}$ with a demand-shifter. Our chosen instrument is the average high temperature in market m at time t, and it squared.[7]

Having estimated the supply equation, the coefficient on occupancy $\widehat{\beta_o}$ gives us an estimate of the price coefficient α according to $\widehat{\alpha} = -1/\beta_o$. Substituting this estimate into the demand equation (1) yields:

$$o_{j,t} - \hat{\alpha} p_{j,t} \equiv o^{adj}_{j,t} = \beta r_{j,t} + \gamma_j + \mu_{m,t} + \xi_{j,t} \tag{5}$$

where we call the new variable $o^{adj}_{j,t} = o_{j,t} - \hat{\alpha} p_{j,t}$ the "adjusted occupancy" of the hotel, as it adjusts the occupancy of a hotel up if it has high prices.

We estimate (5) as our demand equation, by OLS. Unlike the original demand equation (1) we no longer have a price endogeneity concern, since we have used the supply side estimates to put both endogenous variables on the left hand side. This trick eliminates the problem of finding a convincing supply-shifter, which we have been unable to do (candidates we tried and rejected included wages of low-skilled workers, and Hausman-style and BLP-style

[7] The relationship between occupancy and temperature is well approximated by a parabola, peaking between 70 and 80 degrees Farenheit. We have experimented with more flexible specifications, but they don't improve the first-stage very much.

instruments (Hausman, 1996; Berry et al., 1995)).

3.2 Results

Supply and Demand. We present our initial results in the first column of Table 2. All our specifications cluster errors at the hotel level to account for correlation in hotel prices and demand.[8]

Columns (1) to (4) give results for the supply estimation, where (1) and (2) don't include year interactions with reviews, while columns (3) and (4) do. Recall that on the supply side, we instrument for occupancy with temperature and temperature squared. Columns (1) and (3) show that the instruments in the first stage are strong, with an F-statistic of around 4300 in both cases. They also have the expected signs, with occupancy rising in temperature initially and falling at high temperatures.

Columns (2) and (4) of Table 2 present the main supply side results (i.e. second stages). We find either zero or small significant effects of hotel ratings on the supply curve itself, implying that managers do not respond to more favorable reviews by adjusting the supply curve.[9] One reason for this may be that the supply curve is upward sloping i.e. prices go up as the hotel becomes more occupied, so that higher reviews translate into higher prices through increased occupancy. This is shown by the coefficient on occupancy in column (2), which implies that as the hotel goes from empty to full, it increases prices by $71 (or about 69% for a hotel room at an average hotel in our data set). Recall that this coefficient can be interpreted as the inverse of the price coefficient in the demand system, and taking this interpretation implies an average price elasticity of -2.52 for hotels in our dataset.[10]

The last two columns show the results of the demand estimation. In column (5), we a significant effect of review platforms on hotel demand: a one-star increase in a hotel's cross-platform rating is associated with a 0.037 point increase in occupancy. A typical hotel has an average occupancy level of 64%, so this is a 5.8% percent increase in demand.

This estimate captures the average effect of review platforms on hotel demand during our

[8] We do not account for first-stage error in the estimation of $\widehat{\alpha}$ in the demand estimates, as $\widehat{\alpha}$ is quite precisely estimated and so we suspect it would not make much difference. Accounting for this error, e.g. by bootstrapping, would be computationally demanding given the number of estimated fixed effects.

[9] The main exception to this general pattern is the estimated effects of rating on supply in recent years: in 2014, a hotel that increases its rating by 1 rating point would increase its prices by $5.91 — this is arguably still pretty small.

[10] Due to the linear demand specification, the elasticity varies along the demand curve, and so must be evaluated at each individual hotel's price-occupancy pair.

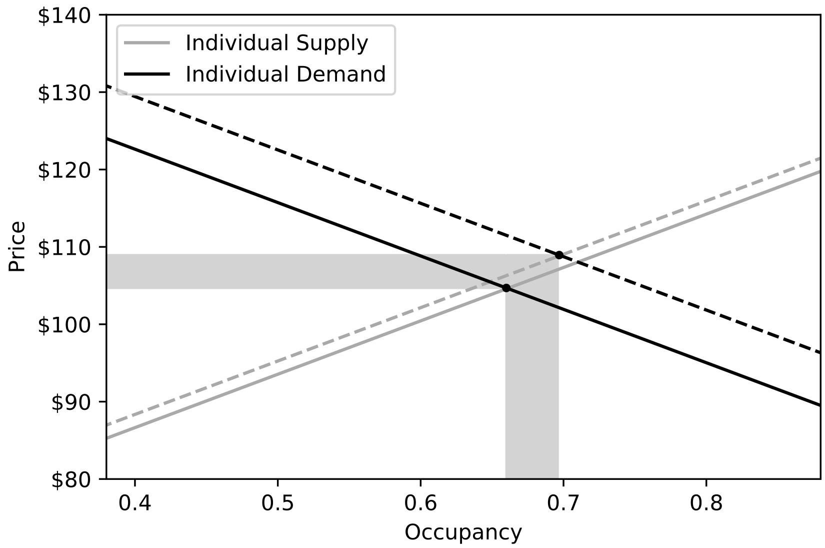


Figure 6: Supply and demand curve shift after a one standard deviation (0.56 stars) increase in average ratings in 2014.

10-year observation period. In column (6) of the Table we relax the assumption that the coefficients are stable over time. As noted earlier, we would expect increasing coefficients, because the informational content of review platforms has increased over time as more reviews have come in. For instance, in 2005, 66% of the hotels in our data had at least one TripAdvisor review. By 2014 this number had increased to 99% of hotels. There has also been increased engagement with these platforms as internet penetration has grown. As expected, we find that review platforms have become more influential over time: by 2014 the impact of a 1-star increase on sales is 0.175 or a 27.8% percent increase in demand. The pattern is not entirely monotone: there is a slight dip in 2009-2010, corresponding to the recession in those years.

We depict our understanding of the market in Figure 6, using the coefficient estimates from the most recent year, 2014, implied by the specifications in columns (4) and (6) of Table 2, for an average hotel in the market (i.e. starting from the average price and occupancy levels at the intersection of the solid lines, moving to a new equilibrium level at the intersection of the dashed lines). The figure illustrates the major forces present in the market: an increase in ratings mostly increases demand (shift from solid to dashed black line) rather than supply (shift from solid to dashed grey line). Still, because the individual supply curve is upward sloping, a 1 standard deviation increase in ratings (an increase of 0.56 ratings points) increases both price and occupancy levels, leading to a revenue increase equal to the shaded light grey area, around 10%.

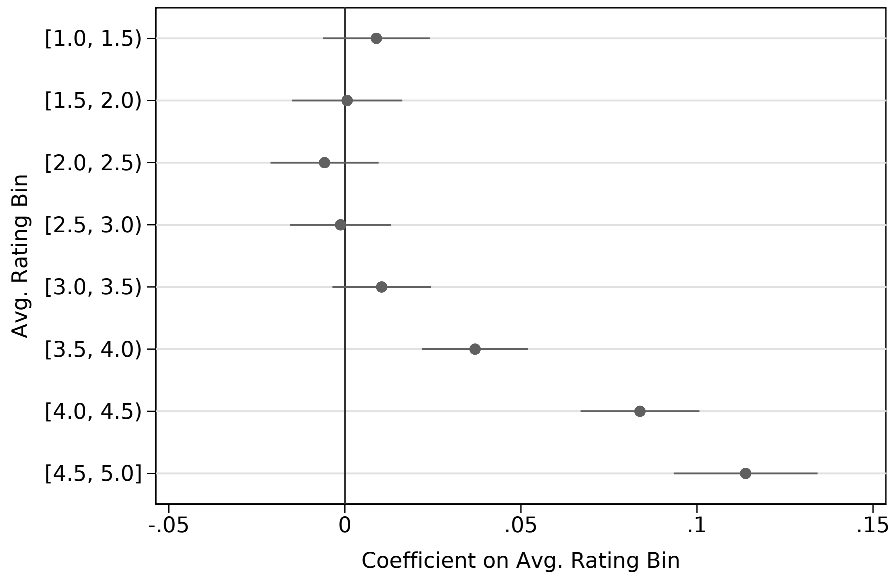


Figure 7: Estimated coefficients of adjusted occupancy on average ratings binned in half-star increments.

Non-Linearity. One might wonder if the impact of reviews is non-linear in rating. To address this question we also run a demand specification similar to equation (5) where the dependent variable is adjusted occupancy, and in which we dummy ratings out by half-rating point instead of including a linear term. The coefficients on the dummies (along with 95% confidence intervals) are plotted in Figure 7. We find that the coefficients are approximately equal to zero for ratings up until around 3.5 stars, after which they increase approximately linearly. This indicates that all hotels with reviews below 3.5 stars (a relatively small fraction of hotels) are treated as equally bad by consumers, but over the relevant range (i.e. most hotels), the linear approximation is good.

Heterogeneity. We next investigate heterogeneity of effects in Table 3, where we interact the treatment variable (i.e. ratings) with a number of observable categories, individually. The first-stage regressions continue to look sensible. On the supply side (columns (2) and (4)), we continue to see small effects of reviews on the supply curve, with the only exception being for luxury hotels.

The interesting results are on the demand side. Column (5) allows reviews to interact with price class (where the order of class ranges from economy through midscale, upper midscale, upscale, upper upscale to luxury). We would expect that reviews matter more at the top, since consumers who can afford to pay high prices are also more discriminating. This is

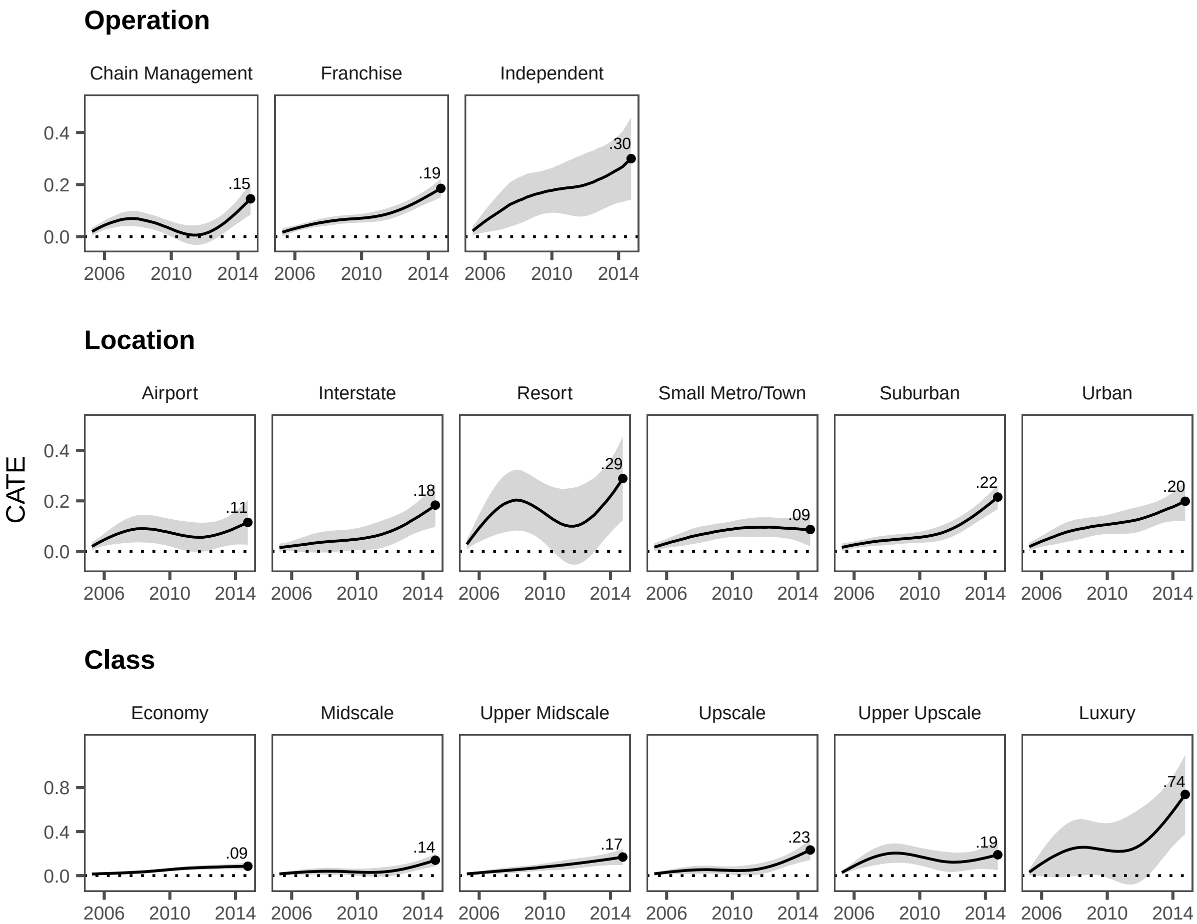


Figure 8: Conditional average treatment effects of ratings on occupancy over time by hotel operation, location, and class (95% confidence intervals shown with grey bands). The annotated dots display effect sizes in the last observation period.

indeed what we find: the estimated coefficients have the expected order, and ratings matter most for luxury hotels.

Column (6) interacts ratings with hotel organizational form (e.g. franchise). Chain and franchise hotels have an informational advantage over independent hotels derived from brand recognition, standardized service, and little variation in quality from location to location. Lacking these advantages, one would expect that high-quality independent hotels should benefit more from an additional information channel for consumers to learn about product quality. We find that the impact of review platforms is substantially larger for independent hotels than chains – 10.4% vs 4.3% for a 1-star increase in rating. This analysis suggests that review platforms have helped narrow the informational gap between chains and independents.

Another way of examining heterogeneity of effects is using some of the new techniques for detecting heterogeneous treatment effects using machine learning (see e.g. Chernozhukov et al. (2017); Athey and Imbens (2016); Athey and Wager (2017); Foster and Syrgkanis (2019); Syrgkanis et al. (2019)).

We implement the double machine learning approach of (Chernozhukov et al., 2017): in a first-stage, we fit the outcomes (sales) and treatments (ratings) flexibly on all of the controls in our dataset, using gradient boosted trees. We use 5-fold cross validation to select the number of trees in each model. Then in a second-stage we run a linear regression of the outcome residuals from that first-stage on the treatment residuals interacted with a subset of the controls, by OLS. We allow for a high-dimensional set of interactions between all dummy variables and a natural cubic spline of the time variable with three equally spaced knots. We use cross-fitting i.e. the residuals for one half of the data are constructed using a model fitted on the other half of the data. This allows for valid inference (otherwise the first-stage errors would contaminate the second stage). The results are shown in Figure 8, which plots the conditional average treatment effects after conditioning on pairs of covariates (always time, and then operation, location and class respectively in the next three panels). In addition to the findings above – that reviews appear to be more important for upscale and independent hotels – we learn that they matter more for suburban and urban hotels, and less for airport and small town hotels, who presumably face less competition. The upward time trend in the importance of reviews is consistent across all types of hotels, but the slope appears to be steepest for luxury and resort hotels.[11]

Impact of Revenue Management. We show how revenue management interacts with ratings in Table 4. The dependent variable in all columns is log revenue. Column (1) shows that revenue increases by 7.6% for a 1 point increase in rating. This estimate of the revenue effect is in line with previous studies (Luca, 2011; Anderson and Magruder, 2012). For example, Luca (2011), considering the impact of a 1 point increase in Yelp reviews for restaurants in Seattle in the period 2003–2009 estimates an effect on revenue of between 5$ and 9%.

Columns (2)-(4) add interactions with the revenue management proxy, for various transformations of the revenue management variable: raw (column (2)), quantile transformation (column (3)), and a dummy for being above or below the median of the revenue management variable (column (4)). In these columns the data used is from a hold-out period, 2010-2014.[12]

[11]Since many resort hotels are luxury hotels, one might expect the time-pattern of the CATES to look similar, as they do.

[12]Recall that we estimated a proxy for the extent to which a hotel uses revenue management techniques

The easiest to interpret is the last column: hotels in the bottom half of the revenue management variable are estimated to increase revenue by 5% after a 1 unit increase in rating, while those in the top are estimated to increase revenue by 12%. This is consistent with a complementarity between investments in revenue and reputation management (though since revenue management is not randomly assigned, we cannot be sure this is a causal relation).

4 Natural Experiment

To further reinforce our causal claims we exploit a natural experiment occurring in our data. In March 2012, Expedia announced its intention to to incorporate reviews from its sister company Hotels.com in its review collection.[13] While Expedia announced these plans in March 2012, the merger did not take place immediately. Unfortunately, we could not find any public announcement from Expedia disclosing the date at which the change was implemented. Instead, we turned to the Internet Archive (IA). The Internet Archive provides access to historic snapshots of billions of web pages going back decades in time. By browsing through IA snapshots of Expedia web pages, we discovered that Expedia started showing Hotels.com reviews in June 2013. We arrived at this conclusion based on two markers: first, during June 2013 Expedia started displaying reviews marked with the phrase "Posted [DATE] on Hotels" (where "[DATE]" is, *e.g.*, "May 30, 2014"); second, between May and June 2013 the reported total review count for each hotel increased abruptly.

The merger affected the ratings displayed on Expedia and Hotels.com. Since Expedia and Hotels.com round their ratings to the nearest 0.1, in many cases where the two initial ratings were similar, the merger had the effect of raising or lowering the rating on one platform by multiples of 0.1, having no effect on the other. For instance, suppose that prior to the change the set of ratings Hotels.com is $H = \{4, 4, 4, 4, 5\}$ and the set of ratings on Expedia is $E = \{4\}$ resulting in rounded ratings of 4.2 and 4 respectively. Post merger (and assuming no new reviews) the rounded rating on Hotels.com remains 4.2 stars, while the Expedia rating increases from 4 to 4.2 stars.

It is this kind of plausibly exogenous variation we exploit in the analysis below. Notice that the rounding is a necessary component for this merger to generate useful variation: the average rating across all three platforms, $r_{j,t}$, is by construction unchanged by the merger

by estimating the slope of their supply curve after conditioning on a wide array of controls, using data from 2005-2009 (see Section 2 for more details).

[13] See `https://viewfinder.expedia.com/news/expedia-overhauls-hotel-reviews-consumers-can-now-sort-verified-reviews-by-shared-interest/`

(the underlying set of reviews remains the same), but the average *rounded* rating $\tilde{r}_{j,t}$ isn't.

Let $r^p_{j,t}$ for $p \in \{E, H, T, EH\}$ be a hotel j's average rating on platform p at time t, and $n^p_{j,t}$ its (cumulative) number of reviews. We will use EH to represent Expedia/Hotels.com post-merger. Additionally, let $\tilde{r}^p_{j,t}$ denote hotel j's *rounded* average rating on platform p.[14] Next, define a hotel's cross-platform volume-weighted average rounded rating as:

$$\tilde{r}^P_{j,t} = \frac{\sum_{p \in P} n^p_{j,t} \tilde{r}^p_{j,t}}{\sum_{p \in P} n^p_{j,t}}. \tag{6}$$

Let $P_1 = \{E, H, T\}$ represent the setting where the review platforms operate independently, and $P_2 = \{EH, T\}$ represent the post-merger setting. Define the difference:

$$\Delta\tilde{r}_{j,t} = \tilde{r}^{P_2}_{j,t} - \tilde{r}^{P_1}_{j,t}, \tag{7}$$

which can be interpreted as the difference between ratings rounded following the merger and the counterfactual rounded ratings had the merger not taken place. Then, our estimating demand equation is:

$$ao_{j,t} = \phi_1 \Delta\tilde{r}_{j,t} \times \text{Post}_t + \gamma_j + \mu_{m,t} + \xi_{j,t} \tag{8}$$

where Post_t is a binary indicator for post-merger time periods. This can be thought of as a difference-in-differences (DD) specification where the "treatment" (a change in average rounded ratings) is continuous and generated by this merger. To consistently estimate a causal effect, we need that treatment assignment is uncorrelated with the error term. This is plausible given that the treatment is generated by a supply-side change (a merger). We estimate this regression using data from a 6-month window on either side of the merger.[15]

We also estimate a supply specification:

$$p_{j,t} = -\frac{1}{\alpha} o_{j,t} + \phi_2 \Delta\tilde{r}_{j,t} + \delta_j + \mu^c_{m,y} + u_{j,t} \tag{9}$$

where again we expect to see that the coefficient ϕ_2 is approximately zero. Because of the simultaneous determination of price and occupancy, we again instrument for occupancy in this regression with temperature and its square.

[14] Recall that Expedia and Hotels.com round ratings to 0.1-star increments whereas TripAdvisor rounds to half-star increments.

[15] To be consistent we use our estimate $\widehat{\alpha}$ from the OLS regressions in generating the adjusted occupancy variable used here. We have also experimented with using the estimated demand slope coming from the supply model estimated only on these 12 months of data. The results are very similar.

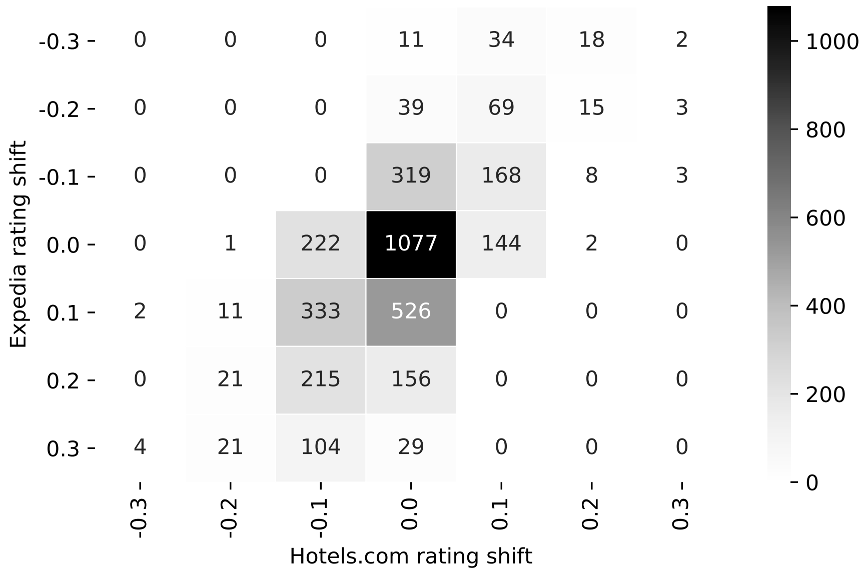


Figure 9: The effect of the Expedia-Hotels.com merger on rounded ratings.

4.1 Results: Natural experiment

We first do some investigation of who is "treated" in our experiment. Figure 9 plots changes in the ratings displayed on Expedia and Hotels.com post-merger, where each cell contains a count of how many hotels are in that cell. Most hotels experience no change, but a fair number increase their rating by an 0.1 increment on one platform while staying fixed or decreasing their rating on the other by 0.1.

The next thing we check is that the treated units are somewhat similar to the untreated on the observables, as one would expect given the quasi-random nature of the rounding. Table 5 shows the characteristics of hotels whose average ratings were affected by the merger (the "rating shift" column), as compared to those who experienced no change. The treatment group, which is larger, comprises hotels who experienced any kind of change, which includes as particular cases having their rating rise on Expedia and fall on Hotels.com, or rise on Expedia and stay constant on Hotels.com (strictly rising or falling on both is mathematically impossible). We see that the treatment and control groups are reasonably balanced on the observables, though treated hotels are statistically significantly younger, cheaper and have more reviews (the last one being a bit of a surprise, since we thought that the ones whose ratings changed might have had fewer reviews). Because of this slight difference between the groups, it is important that we include hotel fixed effects, rather than relying entirely on the quasi-random assignment for identification.

Results are presented in the first three columns of Table 6. The supply side first stage (column (1)) shows that temperature is still a good instrument. The main supply side specification (column (2)) shows that the supply curve is upward sloping, and also that the supply curve shifts up with ratings — though not statistically significantly. Column (3) reports the demand side estimate. We find an effect of ratings on demand that is substantially bigger than those reported in our earlier regressions — approximately a 41% increase in occupancy for a one rating point increase — but less precisely estimated than our OLS results ($p < 0.05$). Our interpretation is that this large positive effect is further evidence of a causal impact of ratings on demand, though we find our earlier and smaller OLS estimates to be more plausible in magnitude.

The last three columns of the table report the results of a Placebo test, where we repeat these regressions, assuming that instead the merger had taken place a year earlier (defining and constructing the variables analogously to how we did in 2013). As one would expect, we find statistical zero coefficients on the ratings shift variable on both the demand and supply side. All in all, the results of the natural experiment, though noisy, suggest that the OLS results we presented earlier in the paper are not substantially upward biased.

5 Conclusion and Implications

It is common wisdom that consumers are devoting increasing attention and time to making informed choices, accessing the many information sources at their disposal. In this paper, we have presented evidence that ratings on popular platforms such as TripAdvisor have causal effects on hotel revenues. This finding is not particularly novel, as others have already documented such effects.

What *is* new is that we have been able to decompose these revenue effects into a supply and a demand component, showing that prices and occupancy increase because the demand curve shifts up, and the supply curve is upward sloping. Revenue management plays a key role here, freeing managers from the concern of adjusting prices in response to fluctuations in reviews by automatically increasing prices as demand grows. Indeed, we find that hotels that are estimated to have better management practices benefit substantially more from an increase in their online reputation that those that don't.

We also provide nuanced findings on how much reviews matter and for whom. Reviews are becoming increasingly important over time, with the coefficients on ratings increasing

33% year on year in recent years. This points to an increased importance of reputation management. This is particularly true for independent hotels (who lack a chain reputation) and luxury hotels (whose customers expect the best).

Our results indicate that consumers rely on review platforms to make better choices, potentially resulting in consumer welfare gains. However, our work also shows that these gains have to be weighed against potential losses in consumer welfare resulting from higher prices at higher-rated hotels. Thus the welfare consequences of review platforms are ambiguous, and estimating them precisely is an interesting direction that we explore in ongoing work (Lewis and Zervas, 2016).

There are other interesting topics for future research. We have not investigated the text content of online reviews, which may play an important role in matching consumers to hotels and is thus another source of surplus. Modeling information acquisition may also be important, as online reviews have almost certainly had an effect on the amount of time that consumers spend acquiring information. Market structure is also presumably affected by review platforms, as the crowd-sourcing of reputation may on the one hand make entry cheaper, but on the other hand increase the market power of high-quality hotels.

References

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

S. Athey and G. Imbens. Recursive partitioning for heterogenous causal effects. *PNAS*, 113 (27):7353–7260, 2016. URL `https://www.pnas.org/content/113/27/7353`.

S. Athey and S. Wager. Estimation and inference of heterogenous treatment efffects using random forests. 2017. URL `https://arxiv.org/abs/1510.04342`.

J. Bai. Melons as lemons: Asymmetric information, consumer learning and seller reputation. 2015.

S. Berry, J. Levinsohn, and A. Parkes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

G. Bitran and R. Caldentey. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–229, 2003.

J. R. Brown, A. Goolsbee, et al. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507, 2002.

E. Brynjolfsson and J. Oh. The attention economy: measuring the value of free digital services on the internet. 2012.

E. Brynjolfsson, Y. Hu, and M. D. Smith. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, 49(11):1580–1596, 2003.

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

Y. Chen and J. Xie. Third-party product review and firm marketing strategy. *Marketing Science*, 24(2):218–240, 2005.

Y. Chen and J. Xie. Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, 54(3):477–491, 2008.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and causal parameters. 2017. URL `https://arxiv.org/abs/1608.00060`.

J. A. Chevalier, Y. Dover, and D. Mayzlin. Channels of impact: User reviews when quality is dynamic and managers respond. *Marketing Science*, 37(5):688–709, 2018.

D. W. Elfenbein, R. Fisman, and B. McManus. Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics*, 7(4):83–108, 2015.

W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management science*, 49(10):1287–1309, 2003.

D. Foster and V. Syrgkanis. Orthogonal statistical learning. 2019. URL `https://arxiv.org/abs/1901.09036`.

A. Fradkin, E. Grewal, D. Holtz, and M. Pearson. Reporting bias and reciprocity in online reviews: Evidence from field experiments on airbnb. 2014. Cited with permission.

A. Goolsbee, P. J. Klenow, et al. Valuing consumer products by the time spent using them: An application to the internet. *American Economic Review*, 96(2):108–113, 2006.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

J. Hausman. *The Economics of New Goods*, chapter Valuation of New Goods under Perfect and Imperfect Competition, pages 207–248. University of Chicago Press, 1996. URL `https://www.nber.org/chapters/c6068.pdf`.

B. Hollenbeck. Online reputation mechanisms and the decreasing value of chain affiliation. *Journal of Marketing Research*, 55(5):636–654, 2018.

B. Hollenbeck, S. Moorthy, and D. Proserpio. Advertising strategy in the presence of reviews: An empirical analysis. *Marketing Science*, Forthcoming.

N. Hu, J. Zhang, and P. A. Pavlou. Overcoming the j-shaped distribution of product reviews. *Communications of the ACM*, 52(10):144–147, 2009.

G. Z. Jin and P. Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, 2003.

D. S. Lee and T. Lemieux. Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355, 2010.

G. Lewis and G. Zervas. The welfare impact of consumer reviews: A case study of the hotel industry. *Working paper*, 2016.

M. Luca. Reviews, reputation, and revenue: The case of Yelp.com. *Harvard Business School NOM Unit Working Paper*, (12-016), 2011.

M. Luca and G. Zervas. Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427, 2016.

D. Mayzlin, Y. Dover, and J. Chevalier. Promotional reviews: An empirical investigation of online review manipulation. *The American Economic Review*, 104(8):2421–2455, 2014.

S. Pantea and B. Martens. The value of the internet for consumers. *Available at SSRN 2446962*, 2014.

D. Proserpio and G. Zervas. Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, 36(5):645–665, 2017.

P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on ebay: A controlled experiment. *Experimental economics*, 9(2):79–101, 2006.

V. Syrgkanis, V. Lei, M. Oprescu, M. Hei, K. Battocchi, and G. Lewis. Machine learning estimation of heterogenous treatment effects with instruments. 2019. URL `https://arxiv.org/abs/1905.10176`.

Y. Wang and A. Chaudhry. When and how managers' responses to online reviews affect subsequent reviews. *Journal of Marketing Research*, 55(2):163–177, 2018.

Table 1: Descriptive statistics.

	Mean	SD	Min	Max
Room Price	103.57	66.50	20.89	1094.27
Occupancy	0.63	0.14	0.01	0.99
Monthly Room Demand	2646.62	3291.00	40.46	74904.25
Monthly Room Supply	3993.53	4266.87	182.50	87035.62
Avg. Rating	3.76	0.66	1.00	5.00
Reviews Per Month	7.49	11.21	0.00	316.61
RM Proxy	25.37	28.97	−95.90	372.11

Table 2: The effect of ratings on occupancy and price.

	(1^{st} stage) Occupancy	Price	(1^{st} stage) Occupancy	Price	Adj. Occ.	Adj. Occ.
Occupancy		70.513***		70.471***		
		(1.549)		(1.549)		
Avg. Rating	0.019***	0.178			0.041***	
	(0.002)	(0.211)			(0.003)	
2005 × Avg. Rating			0.013***	−1.830***		0.000
			(0.002)	(0.275)		(0.004)
2006 × Avg. Rating			0.011***	0.150		0.024***
			(0.002)	(0.275)		(0.004)
2007 × Avg. Rating			0.013***	1.883***		0.053***
			(0.002)	(0.279)		(0.004)
2008 × Avg. Rating			0.014***	2.784***		0.072***
			(0.002)	(0.304)		(0.004)
2009 × Avg. Rating			0.014***	−0.954***		0.019***
			(0.002)	(0.308)		(0.004)
2010 × Avg. Rating			0.023***	−1.489***		0.026***
			(0.003)	(0.340)		(0.004)
2011 × Avg. Rating			0.035***	−0.579		0.062***
			(0.003)	(0.363)		(0.005)
2012 × Avg. Rating			0.041***	0.843**		0.094***
			(0.003)	(0.429)		(0.006)
2013 × Avg. Rating			0.045***	2.845***		0.129***
			(0.003)	(0.527)		(0.007)
2014 × Avg. Rating			0.050***	5.683***		0.180***
			(0.003)	(0.639)		(0.009)
Temp.	0.037***		0.037***			
	(0.001)		(0.001)			
Temp.^{2}	−0.000***		−0.000***			
	(0.000)		(0.000)			
Hotel FE	Yes	Yes	Yes	Yes	Yes	Yes
Market-year FE	Yes	Yes	Yes	Yes	No	No
Market-year-month FE	No	No	No	No	Yes	Yes
F-stat	4294.4		4295.8			
N	442497	442497	442497	442497	442497	442497

Note: Standard errors clustered at the hotel level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Heterogeneous treatment effects of ratings on occupancy and price.

	(1^{st} stage) Occupancy	Price	(1^{st} stage) Occupancy	Price	Adj. Occ.	Adj. Occ.
Occupancy		70.458***		70.499***		
		(1.551)		(1.548)		
Economy × Avg. Rating	0.015***	−1.093***			0.015***	
	(0.003)	(0.299)			(0.003)	
Midscale × Avg. Rating	0.014***	0.122			0.032***	
	(0.004)	(0.424)			(0.006)	
Upper Midscale × Avg. Rating	0.011***	0.308			0.030***	
	(0.003)	(0.416)			(0.006)	
Upscale × Avg. Rating	0.025***	0.674			0.065***	
	(0.005)	(0.591)			(0.010)	
Upper Upscale × Avg. Rating	0.042***	1.290			0.106***	
	(0.006)	(1.215)			(0.018)	
Luxury × Avg. Rating	0.032**	9.096**			0.190***	
	(0.013)	(4.005)			(0.063)	
Chain Management × Avg. Rating			0.018***	−0.276		0.033***
			(0.003)	(0.558)		(0.008)
Franchise × Avg. Rating			0.018***	−0.081		0.038***
			(0.002)	(0.247)		(0.003)
Independent × Avg. Rating			0.020***	2.434***		0.070***
			(0.005)	(0.931)		(0.014)
Temp.	0.037***		0.037***			
	(0.001)		(0.001)			
Temp.^2	−0.000***		−0.000***			
	(0.000)		(0.000)			
Hotel FE	Yes	Yes	Yes	Yes	Yes	Yes
Market-year FE	Yes	Yes	Yes	Yes	No	No
Market-year-month FE	No	No	No	No	Yes	Yes
F-stat	4294.1		4295.6			
N	442497	442497	442497	442497	442497	442497

Note: Standard errors clustered at the hotel level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The effect of ratings on hotel revenue as a function of revenue management.

	log(Revenue)	log(Revenue)	log(Revenue)	log(Revenue)
Avg. Rating	0.083***	0.060***	0.060***	0.061***
	(0.008)	(0.007)	(0.011)	(0.007)
Avg. Rating × RM Proxy		0.001***		
		(0.000)		
Avg. Rating × Normalized RM Proxy			0.045*	
			(0.024)	
Avg. Rating × Binary RM Proxy				0.057***
				(0.017)
Hotel FE	Yes	Yes	Yes	Yes
Market-year-month FE	Yes	Yes	Yes	Yes
N	238510	198135	198135	198135

Note: Standard errors clustered at the hotel level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Platform-merger experiment: balance check for the month prior to the merger.

	Mean Rating shift	Mean No rating shift	Diff.	Std. Error	p-value
Number of Rooms	136.02	131.75	-4.27	4.65	0.358
Hotel Age	26.46	29.43	2.96***	0.65	0.000
Price	105.84	114.16	8.33***	2.50	0.001
Occupancy	0.67	0.68	0.00	0.01	0.576
Monthly room demand	2943.35	2884.07	-59.28	112.76	0.599
Monthly room supply	4223.21	4103.53	-119.68	144.11	0.406
Avg. rating	3.88	3.68	-0.19***	0.02	0.000
Num. reviews at merge	275.01	235.96	-39.05***	13.49	0.004

Note: Hotels grouped by whether they experienced a rating change on at least one platform.

Table 6: Platform merger experiment: the effect of ratings on occupancy and price.

	2013			*2012 (placebo)*		
	(1^{st} stage) Occupancy	Price	Adj. Occ.	(1^{st} stage) Occupancy	Price	Adj. Occ.
Occupancy		79.035^{***}			69.497^{***}	
		(1.840)			(1.691)	
Rating shift	0.093	12.041	0.406^{**}	−0.010	−2.068	0.051
	(0.060)	(9.648)	(0.201)	(0.062)	(8.729)	(0.191)
Temp.	0.037^{***}			0.037^{***}		
	(0.001)			(0.001)		
Temp.^{2}	-0.000^{***}			-0.000^{***}		
	(0.000)			(0.000)		
Hotel FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	No	No	Yes
F-stat	3773.4			3438.7		
N	50,708	50,708	50,708	49,487	49,487	49,487

Note: Standard errors clustered at the hotel level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$