The Welfare Impact of Consumer Reviews: A Case Study of the Hotel Industry[†]

Gregory Lewis
Microsoft Research and NBER

Georgios Zervas
Questrom School of Business
Boston University

[Preliminary and incomplete — please do not cite]

July 13, 2016

Abstract

Platforms such as Yelp and TripAdvisor aggregate crowd-sourced information about users' experiences with products and services. We analyze their impact 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. Both hotel demand and prices are positively correlated with their average ratings on TripAdvisor, Expedia and Hotels.com, and such correlations have grown over our sample period from a statistical zero in the base year to a substantial level today: a hotel rated one star higher on all the platforms on average has 27% higher demand, and charges 10% more. 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 a structural model of supply and demand with partially informed consumers, finding that in a counterfactual world without any review information, consumers surplus would fall by 65¢ per room-night ($1.1bn when extrapolating to the US hotel industry as a whole).

[†]We thank Michael Kurish for excellent research assistance. We also thank STR, and in particular Duane Vinson, for helping us with data acquisition.

1 Introduction

The internet has been one of the most significant innovations of recent times. It has made information about goods and services more readily available, allowing consumers to make more informed choices. Online review platforms are one important source of information. For example, TripAdvisor — the world's largest travel review platform — contains 320 million reviews. Yelp, which aggregates reviews for local businesses receives approximately 150 million visitors per month.[1] But the welfare implications and magnitudes of such information access are unclear. For example, when everyone in New York knows what the top ramen shops are, these businesses may raise prices or have long queues to ration demand, which in turn may actually leave consumers worse off.

We investigate how the increasing popularity of these information intermediaries has affected consumer choice, and how firms have responded to these changes in consumer behavior. The specific setting we investigate is that of the hotel industry. Given how the internet has transformed travel planning and booking, this an appealing setting to study.

Specifically, we ask three main questions. First, to what extent are information intermediaries helping consumers make better choices? Second, how are firms responding? Third, what are the welfare consequences of these forces combined? Our focus is on the market as a whole, rather than the effects of increased ratings on individual firms.

Our analysis is enabled by the combination of two novel data sources. The first data source we employ is a decade-long proprietary panel of hotel financial performance. The panel, which is at the month level, contains information on the revenue and demand (measured in room-nights) of roughly one half of all hotels that operated during our observation period in California, Oregon, Washington, Nevada, and Arizona. Our second source of data is the entire historical record of online reviews for these hotels, which we collect from three major review platforms: TripAdvisor, Expedia (the world's largest online travel agent), and Hotels.com (an Expedia brand).

A unique feature of our data which allows us to identify the impact of information disclosure on consumer choice, is that it spans a period during which review platforms grew from insignificant sources of information to the major source of travel information. This allows us to analyze long-run trends. For example, one measure of the quality of a hotel is its average rating (across all three platforms) as of December 2014, when most hotels have a substantial number of reviews. Measured this way, the average quality of hotels that consumers stay at has been growing modestly over time, from around 3.8 to 4.0. This suggests

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

that the impact of these platforms on consumer choice has been correspondingly modest: long before TripAdvisor and Expedia became popular information sources, consumers tended to choose hotels that are only of slightly lower *current* ratings than the ones they choose today.

Of course, there are other potential explanations, such as higher-quality hotels raising their prices, so that they are chosen no more often than they were in the past. So to analyze this more carefully, we estimate a logit demand system that controls for price, time-invariant hotel characteristics (through hotel fixed effects) and market-specific demand shocks (though market-year-month fixed effects). As a discrete choice model, it is also able to account for changes in competition arising from the information available to consumers about competing hotels. A maintained assumption of this model is that all consumers account for the review platform information when making choices. The reality is surely more complicated: more tech-savvy consumers and those with a strong taste for hotel quality are more likely to seek out this information. We think of our model as in some sense a "linearizaton" of a more complex mixture model.

Our key findings are: 1) by the end of our observation period a 1-star increase in a hotel's rating is associated with a 27% increase in a hotel's market share, and 2) hotels respond to a 1-star increase by raising prices approximately by 10%. This suggests that hotels are able to capture some of the welfare generated by review platforms. We also show that the *difference* between the average TripAdvisor and Expedia rating is significantly correlated with demand, negatively towards the beginning of our sample period (when Expedia was significantly more well-known), and positively more recently (when TripAdvisor's growth in internet traffic has been much quicker than Expedia's). While these analyses point towards a causal role for review platforms, they do not entirely isolate the impact of crowdsourced reviews from other information sources consumers might have access to.

To disentangle the causal impact of review platforms from other confounding factors, we exploit a natural experiment in our data. In June 2013, Expedia and Hotels.com merged their review collections, so that a hotel with different average ratings on the two sites prior to the merger displayed the same average rating after the merger. Because of the way ratings on these platforms are rounded to the nearest 0.1, it is often the case that a hotel's rating goes up by 0.1 on one platform, while remaining unchanged on the other, generating variation in the average displayed rating of the hotel. Treating the merger as exogenous, we estimate changes in hotel demand around around the date of the natural experiment using a difference-in-differences (DD) empirical strategy. Specifically, by comparing changes in demand for hotels whose ratings changed against hotels whose ratings were unaffected, we find that a 1-star increase is associated with a 26-29% increase in demand. These estimates

are strikingly similar to our previous analyses.

The size of the treatment effect size we estimate may seem surprisingly large, but becomes more plausible when evaluated against the effort hotels would have to expend to achieve a 1-star increase and realize the associated gains in demand. For instance, consider a hotel with a 4-star average rating and 100 reviews, which resembles the median hotel in our data. Such a hotel would need an uninterrupted streak of two thousand 5-star ratings to increase its average rating by 1 star. Since most changes in our data are much smaller that 1-star, so are the changes they cause in hotel demand.

In the last part of the paper, we estimate a counterfactual in which the information available on the review platforms disappears and consumers assume that all hotels in a market are of equal (latent) quality (they still differentiate hotels based on observables). We estimate how this affects choice, consumer surplus and revenue. We find that counterfactual consumer surplus is between 63¢ and 66¢ lower, per room night, depending on whether or not hotels are allowed to respond with price adjustments. Though this estimate seems small compared to the price of an average hotel stay (around $100 per night), it still adds up to a substantial amount when aggregated over all the room nights booked in the US as a whole (over $1.1bn).[2]. This counterfactual estimate does not take into account changes in the costs of consumer information acquisition, which are not modeled here.

The impact of reviews and ratings on firms' sales has been a popular subject in the marketing and economics literatures. For example, 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 (Chevalier and Mayzlin, 2006; Cabral and Hortacsu, 2010; Luca, 2011; Anderson and Magruder, 2012). These papers 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 reputation is a significant driver of sales. Our work differs from these papers in that we take a market-level view rather than a firm-centric one, accounting for how changes in information affect competition between firms within a market. In addition to estimating the impact of better ratings on the bottom line of firms, we are able to measure the welfare implications for consumers.

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

[2]According to the American Hotel and Lodging Association, 4.8 million room nights are sold each day in the US, or 1.75bn a year, implying $1.138bn in surplus (source: `https://www.ahla.com/information/`).

on consumer welfare. Elfenbein et al. (2014) 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.

Finally, our work is directly related to the literature on measuring consumer surplus from the digital economomy (Pantea and Martens, 2014). A difficulty in calculating the internet's consumer surplus is that much of what the internet provides is free to consumers. Therefore, it is difficult to measure surplus through the usual means of estimating demand elasticities for online services. To overcome this difficulty, (Goolsbee et al., 2006) measure the consumer surplus of the internet by relating it to the opportunity cost of pursuing non-internet activities. A recent study by (Brynjolfsson and Oh, 2012) relies on similar methods. Both studies find significant gains associated with internet use, in the order thousands of dollars per consumer per year. In this paper, we take a different approach to measuring consumer surplus: while access to review platforms is free, consumers who consult them can make better choices for goods and services the purchase. By observing realized choices and comparing them against choice consumers would have made in the absence of review platforms, we can directly estimate the economic gains of online information intermediaries. Our work also contributes to our understanding of the specific channels through which internet use benefits consumers. The literature has primarily focused on benefits derived from decreased search costs and easier comparison shopping, which in turn have resulted in increased competition among firms, lower prices (Brown et al., 2002), and richer product assortments (Brynjolfsson et al., 2003). In this paper, we show that information sharing among consumers on crowdsourced review platforms is another digital channel through which consumers have realized significant gains.

2 Theory

We begin the paper with a simple model of how information intermediaries such as TripAdvisor and Expedia can influence supply and demand in the hotel market. Consider a Hotelling model with two hotels A and B located at opposite ends of a unit interval. We assume linear transportation costs normalized to one, so that a consumer located at $t \in [0, 1]$ earns utility $u_{A,t} = v_A - p_A - t$ from staying at hotel A and $u_{B,t} = v_B - p_B - (1-t)$ from staying at hotel B. Hotels can be either high-quality ($v_j = H$) or low-quality ($v_j = L$), for $j = A, B$ and $H > L$. Consumers have a prior that high and low-quality are equally likely, and therefore in the absence of further information, have an expected gross utility (i.e. gross of price and

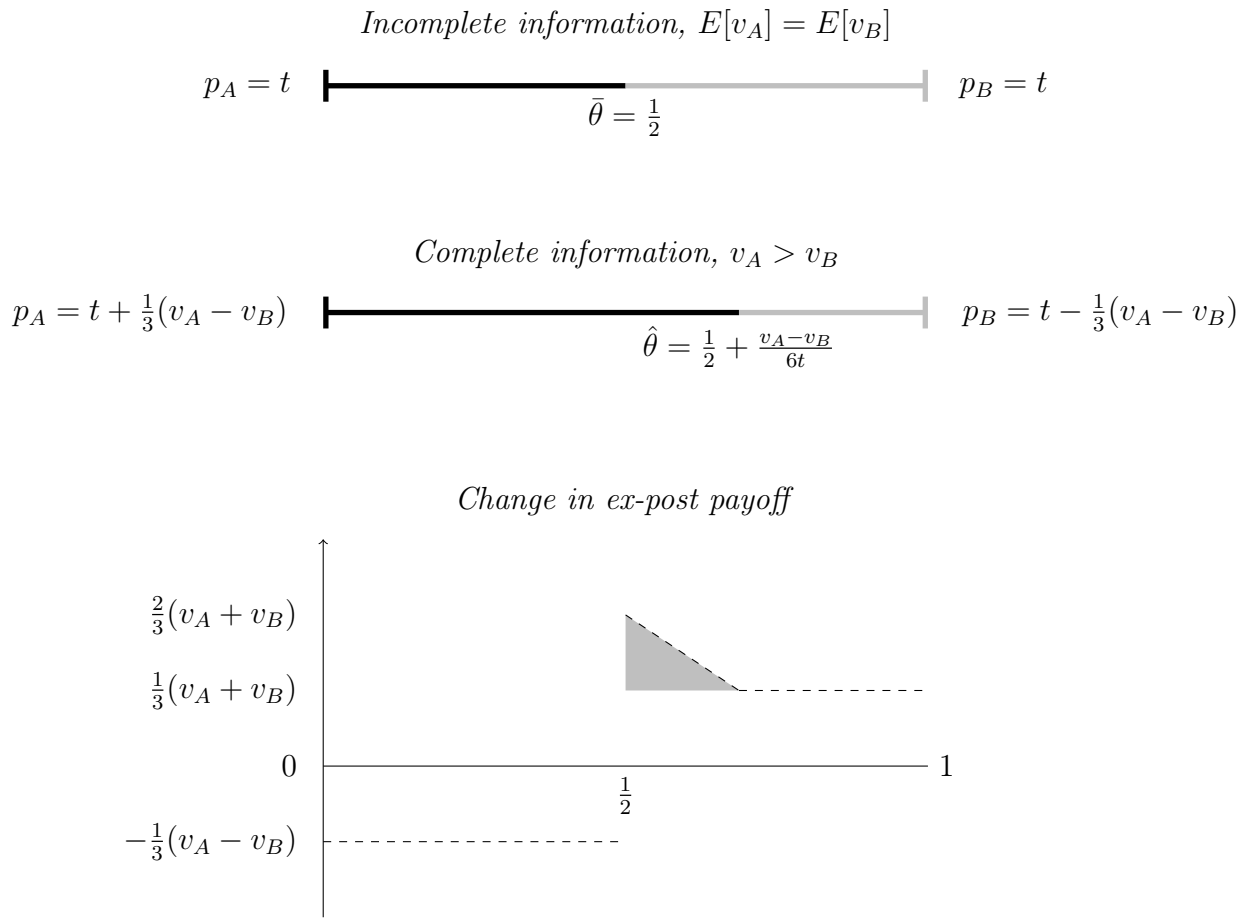


Figure 1: Welfare effects of information under a Hotelling model with linear transportation costs. Firm A is high-quality, and firm B is low-quality. Firm A's market share is drawn in black, and firm B's in grey. Under incomplete information (top panel) consumers expect the firms to provide equal quality. Hence, the indifferent consumer is located at $\bar{\theta} = 1/2$ and firms set equal prices and split the market evenly. Under complete information (middle panel) consumers know that $v_A > v_B$. The indifferent consumer is now located at $\hat{\theta} > \bar{\theta}$. Firm A increases prices and captures a greater market share. The bottom panel displays changes in consumer welfare between the two information schemes as a function of consumer location.

transportation costs) of $1/2(H+L)$ for each hotel. To complete the model, we assume that consumers are uniformly distributed on $[0,1]$ and the hotels have zero marginal costs.

Information intermediaries affect market outcomes by making hotel quality known to consumers through reviews and ratings. To understand how this works, we consider two scenarios: one in which consumers are uninformed about hotel quality (and therefore make decisions based on their prior), and one in which they are perfectly informed. In the perfectly informed case, we will analyze a particular realization in which hotel A is high-quality ($v_A = H$) and hotel B is low-quality ($v_B = L$).

Figure 1 depicts the equilibria in these two different cases. We first consider the incomplete information case. The indifferent consumer is located at

$$\bar{\theta} = \frac{p_B - p_A + t}{2t}$$

and demand for the two hotels is $D_A = \bar{\theta}$ and $D_B = 1 - \bar{\theta}$. From the first order conditions of the firms' profit maximization problem it is easy to see that $p_A = p_B$ and hence $\bar{\theta} = 1/2$, *i.e.*, the firms split the market evenly as shown in the top panel of Figure 1.

We next consider the case where information intermediation permits the two hotels to differentiate by quality (middle panel of Figure 1). Consumers know that $v_A > v_B$ and the location of the indifferent consumer is now given by

$$\hat{\theta} = \frac{p_B - p_A + v_A - v_B + t}{2t}.$$

From the first order conditions of the firms' profit maximization problem it follows that optimal prices are $p_A = t + (v_A - v_B)/3$ and $p_B = t - (v_A - v_B)/3$. Under these prices we have that $\hat{\theta} = 1/2 + ((v_A - v_B)/6t) > \bar{\theta} = 1/2$. Therefore, in this example, information intermediation benefits higher-than-average quality firms both through higher prices and increased market shares (and conversely hurts lower-than-average quality firms).

For consumers, the effects of intermediation intermediation are more ambiguous. The bottom panel of Figure 1 plots changes in consumer welfare as a function of consumer location. First, we consider consumers located in $[0, 1/2)$, whose choice of hotel A is unaffected by information intermediation. Quality disclosure results in these consumers paying a higher price for the same hotel, while their transportation costs and the quality levels they experience remain the same. Therefore, information intermediation leaves consumers closer to firm A than firm B strictly worse off. Next, we consider consumers located in $(\hat{\theta}, 1]$, whose choice of hotel B also remains unaffected. Quality disclosure strictly benefits these consumers because it leads their chosen hotel to reduce prices. Due to the symmetric price adjustments

of the two firms, the welfare gains of a consumer who continues choosing B are equal to the welfare losses of a consumer who continues choosing A, as shown by the dashed line in the bottom panel of Figure 1. Finally, we consider *switchers*, *i.e.*, consumers located in $[\bar{\theta}, \hat{\theta}]$ who switched from the low-quality to the high-quality hotel because of information intermediation. The benefits to switchers vary by location. Information intermediation produces the largest welfare gains for consumers who are equidistant to the two firms. The welfare difference of switching from hotel B to hotel A for a consumer located $\bar{\theta} = 1/2$ is $v_A - v_B$ while the difference in price is $1/3(v_A - v_B)$ for a net utility increase of $2/3(v_A - v_B)$. As the distance to the high-quality hotel increases the welfare gains of switching hotels decrease. The shaded triangle represents the aggregate increase in consumer welfare from information intermediation.

Notice that this welfare analysis is all ex-post (i.e for a particular realization of the quality levels of the two hotels). One could repeat this analysis for each of the remaining possible realizations (H, H), (L, H) and (L, L). The case (L, H) is identical to the present one up to labeling, and yields the same conclusion: not all consumers benefit from information intermediation, and it is the switchers who benefit most. When the hotels have equal ex-post quality, information intermediation has *no real effects*: prices are unchanged, consumers make the same choices, and welfare is unchanged. Intuitively, this is because the prior and the realization agree in their assessment that the two hotels offer equal quality.

We can also compute the ex-ante welfare consequences of information intermediation (i.e. taking the average over all possible realizations). Firms gain $\frac{1}{4t}\left((v_A - v_B)/3\right)^2$, because with some probability they will have increased market power, and the returns to market power are convex in this model. Consumers who never switch (i.e. those outside of $[1 - \hat{\theta}, \hat{\theta}]$) receive the same expected consumer surplus, as the price changes cancel out when taking the expectation. The more elastic consumers in the center receive strictly higher consumer surplus, as when they optimally choose to switch from the the close low-quality hotel to the farther away high-quality hotel, they experience a net utility gain.

This simple static theory foreshadows our main results. As far as firms are concerned, the anticipated impact of information intermediation is clear: higher quality firms should increase prices and capture greater market shares. The welfare gains for consumers depend on how much they value quality relative to other characteristics (in this model, location). Those for whom quality is relatively unimportant will not change their decisions, and in expectation experience no change in consumer surplus. But consumers who value quality highly will benefit from information intermediation, even though firms with high-quality products raise prices. The aggregate change in consumer surplus thus depends on the distribution of taste for quality.

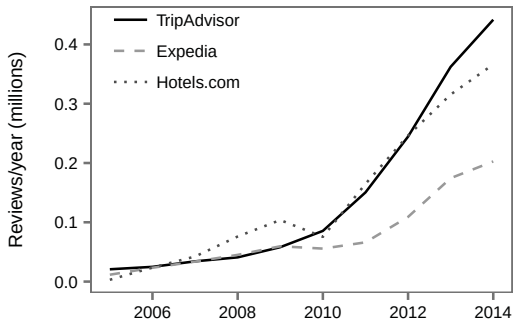


Figure 2: Number of reviews submitted by year.

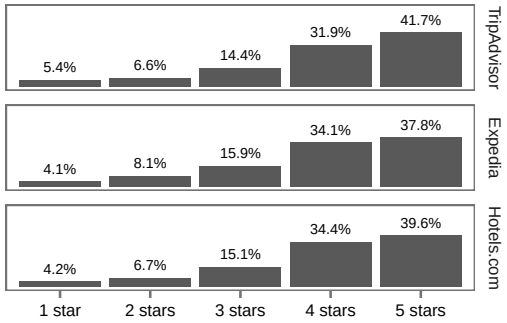


Figure 3: Distribution of ratings by review platform.

Our simple model of information intermediation also yields a few conclusions that inform our empirical analysis. First, to measure the change in consumer surplus from information intermediation, we will need to measure the demand for hotels of different quality levels (whose empirical counterpart is average ratings). As a practical matter, the quality and attention paid to platforms like TripAdvisor may change over time, and so we will want to allow the demand for high-rated hotels to be time-varying. Given such estimates, we will need to infer the demand that we would obtain in a counterfactual world without information intermediation. This in turn will require making some assumption about the prior beliefs of consumers. Second, we will be unable to correctly measure the change in consumer surplus, revenues, and welfare due to information intermediaries without accounting for price responses. We next turn to the data we will use to measure these objects.

3 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 5,944 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: 90% of chains and 10% of independents report data to STR. 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), organi-

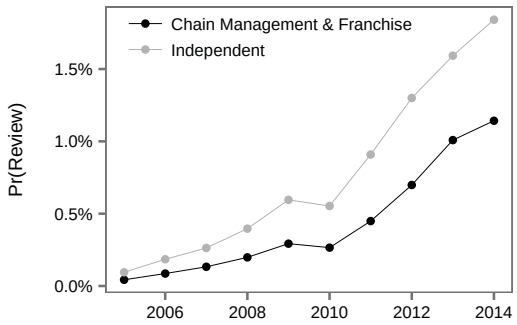


Figure 4: Probability of leaving a review on TripAdvisor, Expedia, or Hotels.com by year.

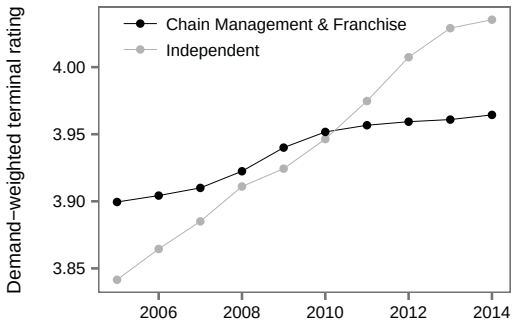


Figure 5: Demand-weighted terminal rating by year.

zational form (chain, franchise, or independent), capacity, and the square footage of any meeting and business facilities. Our data masks the identities of individual hotels.

We augment the hotel financial performance data with a panel of consumer reviews from three major review platforms: TripAdvisor, Expedia, and Hotels.com. Figure 2 plots the (non-cumulative) number of reviews submitted to each the three platforms by year. 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 at the hotel-month level to match the financial performance data. Figure 3 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).

3.1 Descriptive evidence

While review platforms have been growing in popularity little is known about the frequency with which consumers decide to leave a review. Figure 4 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.[3] We find that, while

[3]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

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).

While only a small fraction of consumers leave reviews, a much larger fraction consult them when making choices. Our interest in this paper centers in understanding how consumer choice has responded to the proliferation of online reviews. Here, we provide some descriptive evidence that as review platforms have grown in popularity consumers have made better choices. The quantity we analyze to demonstrate this effect is a hotel's terminal rating, which we define for each hotel as its average rating computed using reviews on all three review platforms by the end of 2014.[4] We think of a hotel's terminal rating as the best proxy for quality a consumer could have had when choosing a hotel. Figure 5 displays the year-by-year evolution of demand-weighted terminal ratings for chain and independent hotels. We observe that as review platform adoption grows, travelers appear to be making better choices, at least as measured by terminal hotel ratings. Interestingly, we observe a steeper trend for independent hotels. This observation coincides with our current understanding of organizational form and reviews. For instance, Luca (2011) shows that chains benefit less from reviews, and Mayzlin et al. (2014) show that chains are less likely to commit review fraud, in part because they benefit less from better reviews. Overall, chains accrue smaller informational gains from online reviews because they provide standardized services and benefit from brand recognition. Therefore, a user reading online reviews will on average acquire less new information about a chain compared to an independent hotel.

4 Empirical strategy and results

We divide our econometric interrogation of the data into two parts: first, we examine how a hotel's ratings affect its demand, and second how its ratings affect the prices it charges. To analyze demand we will need a demand system with partially informed consumers. We offer a simple modification of the discrete-choice logit demand system in which consumers form beliefs about latent hotel quality using all the information at their disposal, and maximize expected utility. In the demand specification, in addition to the usual price endogeneity concerns, we need to address the potential endogeneity of ratings in order to plausibly identify a causal effect. For the most part, our strategy is simply to include flexible controls

`https://www.ahla.com/content.aspx?id=36332`.

[4]To be clear, the average rating is computed as the sum of the rating received by a hotel across all three platforms, divided by the total number of reviews across the platforms.

in order to limit such concerns. In 4.3 we disaggregate our analysis and estimate platform and time specific coefficients (e.g. the coefficient on Tripadvisor ratings in January 2007). We show that these coefficients track engagement with the various platforms (as measured by Google Trends) closely, consistent with a causal interpretation of our earlier regressions. In subsection 4.4 we go further, exploiting a natural experiment based on a merger of the review databases of Expedia and Hotels.com to estimate a causal effect.

4.1 The impact of review platforms on hotel demand

We estimate a logit demand system with imperfectly informed consumers. Consumer i gets utility from choosing hotel j in market m at time t given by:

$$u_{i,j,t} = x_{j,t}\beta + \alpha p_{j,t} + \gamma_{j,t} + \xi_{j,t} + \varepsilon_{i,j,t}$$

where $x_{j,t}$ are hotel characteristics (observed by both agents and the econometrician), $p_{j,t}$ is the price, $\gamma_{j,t}$ is a hotel characteristic unobserved both by the agent and the econometrician, $\xi_{j,t}$ is a hotel characteristic observed by the agent but unobserved by the econometrician, and $\varepsilon_{i,j,t}$ is distributed iid extreme value type 1.

Consumers maximize expected utility:

$$E[u_{i,j,t}|x_{j,t}, p_{j,t}, r_{j,t}, \xi_{j,t}, \varepsilon_{i,j,t}] = x_{j,t}\beta - \alpha p_{j,t} + E[\gamma_{j,t}|r_{j,t}] + \xi_{j,t} + \varepsilon_{i,j,t}$$

for $r_{j,t}$ the average rating of hotel j at time t. We assume the average rating is *sufficient* for the unobserved hotel characteristic $\gamma_{j,t}$, or at least the consumer believes it to be so — they do not condition on $x_{j,t}$ or $p_{j,t}$ in forming their beliefs.

Assume also that the conditional expectation is linear, so $E[\gamma_{j,t}|r_{j,t}] = \phi r_{j,t}$. Then denoting the mean utility (before realization of the idiosyncratic error $\varepsilon_{i,j,t}$) as $\delta_{j,t}$, we get a linear specification:

$$\delta_{j,t} = x_{j,t}\beta + \alpha p_{j,t} + \phi r_{j,t} + \xi_{j,t} \tag{1}$$

As is well known, in logit demand systems, the mean utility index $\delta_{j,t}$ is equal to $\log\left(\frac{s_{j,t}}{s_{0,m(j),t}}\right)$, where $s_{j,t}$ is the market-share of hotel j in time period t. The market-share is in turn defined as $\frac{q_{j,t}}{M_{j,t}}$ where $q_{j,t}$ is the number of room-nights hotel j sold in period t and $M_{j,t}$ is the number of potential buyers in market j at time t. The term $s_{0,m(j),t} \equiv \left(1 - \frac{\sum_{j \in M} s_{j,t}}{M_{j,t}}\right)$ is the share of consumers choosing the "outside option" of non-purchase in the market $m(j)$ that hotel j is

operating in. This gives us a potential estimating equation:

$$\log\left(\frac{s_{j,t}}{s_{0,m(j),t}}\right) = x_{j,t}\beta + \alpha p_{j,t} + \phi r_{j,t} + \xi_{j,t}$$

where all the variables are observed except for the error term $\xi_{j,t}$.

In practice, we opt to estimate something that is similar and has the same structural interpretation, but is more flexible:

$$\log q_{j,t} = \phi_1 \left(r_{j,t} \times v_{j,t}\right) + \phi_2 v_{j,t} + \alpha p_{j,t} + h_j + (\tau_t \times m_j) + \xi_{j,t}, \tag{2}$$

On the LHS, we have replaced the log-share-ratio with the log quantity of rooms sold, avoiding making an assumption as to the unobserved market size $M_{j,t}$. To account for the fact that we have omitted dividing by the market size, we add market-year-month fixed effects $\tau_t \times m_j$ which flexibly control for unobserved demand shocks (such as seasonal demand variation and time trends) contemporaneously affecting all hotels within a specific market. We replace time-invariant hotel characteristics with a hotel fixed effect h_j to flexibly control for them. The average rating $r_{j,t}$ is the average review across TripAdvisor, Expedia and Hotels.com. As a practical matter, not all hotels are rated at all points in time, and so to deal with this we include a dummy $v_{j,t}$ indicating whether the hotel had been reviewed by period t, and add this as an additional control.

One problem with directly estimating Equation 2 is that prices are endogenously determined. This is a more challenging problem than usual due to modern revenue management techniques: demand shocks are quickly transmitted to prices because hotels increase the price of their rooms as they sell out. Therefore, to consistently estimate the equation we would like to instrument for price. Unfortunately, good price instruments are hard to come by in the hotel market. Most variation in prices is driven by correlated demand shocks, implying that Hausman-style instruments are inappropriate. Since we allow for market-year-month fixed effects, valid cost instruments would have to shift individual hotel marginal costs within a market rather than marginal costs as a whole, and such instruments are difficult to find. BLP instruments based on the average characteristics of competitors do not vary much in our data, especially at the market-year-month level. We are not the first to encounter this problem. Koulayev (2014) also seeks hotel price instruments but settles for a rich set of controls instead, as he is met with the same problems we discover here.

Instead, we estimate the price coefficient α by enforcing a set of supply-side moments that require that marginal revenue, assessed at the prevailing prices, should be equal across the high season and low season (where high season and low season are defined at the state

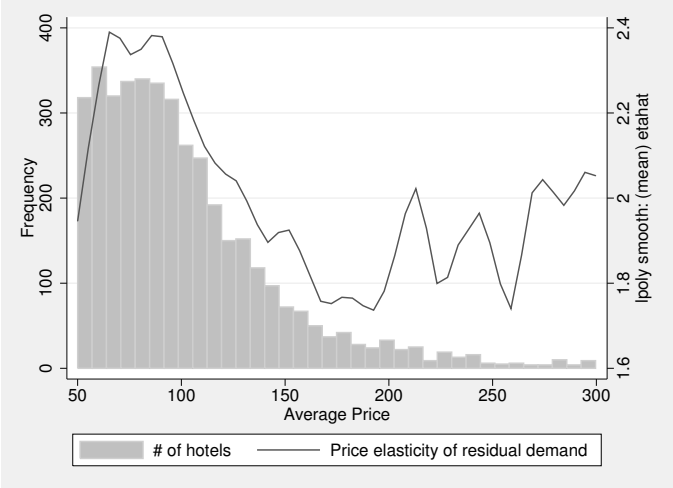


Figure 6: Price elasticity estimates.

level). These follow from price optimization by the hotel (plausible given the aforementioned revenue management) and an assumption that marginal costs are constant over the sample period. Using a flexible semi-parametric demand specification, we recover estimates of the price elasticity of residual demand at each point in a grid of average prices and quantities.

Figure 6 illustrates the results of this procedure. The histogram shows the frequency of hotels whose average price over the sample period falls into each range. The distribution is clearly skewed, with the median hotel charging around $100 but a tail of hotels charging much higher prices. Superimposed on this is a local polynomial plot of the estimated price elasticity of residual demand. Except in the tails, where the estimates become noisy (a variable bandwidth may be more appropriate), the estimated residual demand elasticity tracks the frequency of hotels quite well. This makes sense since residual demand will depend in part on the number of close competitors, and a reasonable notion of closeness is given by competitors who charge similar prices and are of similar size. Moreover, the absolute levels of the estimated elasticities are reasonable, ranging from -1.7 to -2.4 (the y-axis plots the absolute value).

In a logit demand system, the elasticities take the form $\eta_{j,t} = \alpha p_{j,t}(1 - s_{j,t})$ and so we project the estimated elasticities $\hat{\eta}_j$ from this procedure onto $p_{j,t}(1 - s_{j,t})$ to get an estimate $\hat{\alpha} = -0.018$, implying an average elasticity of around -1.8 (since average prices are $\approx$ \$100 and shares are small). Further details are given in the appendix. Given our estimate $\hat{\alpha}$, we re-arrange Equation 2 as:

$$Q_{j,t} = \phi_1 r_{j,t} \times v_{j,t} + \phi_2 v_{j,t} + h_j + \tau_t \times m_j + \xi_{j,t}, \tag{3}$$

where $Q_{j,t} = \log q_{j,t} - \hat{\alpha} p_{j,t}$. We will refer to $Q_{j,t}$ as the *adjusted log quantity*. Equation 3 is the main specification we estimate. The coefficient of interest is ϕ_1, which we interpret as the impact of review platforms on demand. We present our initial results in the first column of Table 1. All our specifications double cluster errors at the market-year and hotel level to account for correlation in hotel prices and demand. We find a significant effect of review platforms on hotel demand: a one-star increase in a hotel's cross-platform rating is associated with a 6% increase in room-nights sold. The effect size we estimate is in line with previously reported estimates (Luca, 2011; Anderson and Magruder, 2012).

This estimate captures the average effect of review platforms on hotel demand during our 10-year observation period. However, the popularity of review platforms has increased dramatically over the past few years. 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.

This motivates our next analysis, where we estimate the impact of review platform over time. To do this we interact $r_{j,t} \times v_{j,t}$ and $v_{j,t}$ with year dummies. We present these results in the second column of Table 1. 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 27%.

Our next set of regressions is motivated by the observation that independent hotels should benefit more from review platforms. Chain 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.

Chain hotels can be operated by the parent company, or they can be franchised. This organizational form distinction is not disclosed on review platforms, and is typically not known to consumers. For this reason, we do not distinguish between these two organizational forms, and simply refer to both of them as *chains*.

To test this hypothesis, we interact $r_{j,t} \times v_{j,t}$ and $v_{j,t}$ with a dummy for independent hotels. We present the results of this regression in the third column of Table 1. We find that the impact of review platforms is substantially larger for independent hotels than chains – 11.5% vs 5.9% for a 1-star increase in rating. This analysis suggests that review platforms have helped narrow the informational gap between chains and independents.

The results we have presented in this section estimate the impact of changing a hotel's rating assuming no change in prices (since price is controlled for). In practice, hotels could adjust their prices in response to quality disclosure to maximize revenue. Next, we investigate hotels' price responses.

4.2 The impact of review platforms on hotel prices

As our simple theory of Section 2 argues, when hotel quality is disclosed, higher quality hotels should respond by raising prices as long as they have some market power. Similarly, lower quality hotels should respond by lowering prices. This response allows hotels to capture some of the welfare that review platforms generate. The decision to raise prices needn't be "active" in the sense that hotel managers actually track online ratings and respond by making decisions; "passive" revenue management systems that charge successively higher prices as each "bucket" of rooms sells out will achieve the same effect.

Our next set of regressions investigates this hypothesis. The specification we estimate takes the following form:

$$p_{j,t} = \phi_1 r_{j,t} \times v_{j,t} + \phi_2 v_{j,t} + h_j + \tau_t \times m_j + \varepsilon_{j,t}, \tag{4}$$

which is similar to Equation 3, except that now the dependent variable we analyze is hotel prices. We present our results in the first column of Table 2. We find that a 1-star increase in a hotel's cross-platform rating is associated with $10 increase in price for a room-night. Given that the average cost of room night in our data is about $100 this represents roughly a 10% increase. This result suggests that hotels are capturing some of the welfare generated by review platforms – even though more consumers are on average making better choices, these choices tend to cost more than they would in the absence of review platforms.

Following our previous steps, we also estimate a model that includes interaction with year dummies to measure changes in price responses over time. We present these results in the second column of Table 2. We find a pattern of increasing prices responses, consistent with our earlier observation that the demand response to reviews is increasing over time.

Last, we estimate a model with an interaction for independent hotels. Our results, as displayed in the third column of Table 2, indicate the independent hotels have been able to raise their price by $3 more per room night than chains. Again, this result is consistent with independent hotels being at an larger informational disadvantage prior to the adoption of review platforms.

4.3 Evidence on causality from trends in platform engagement

We would like to argue that the estimated effects of reviews on demand and prices are causal, but some caution is warranted. Any time-varying omitted demand shock that is correlated with hotel reviews will bias the review coefficients. A simple story in favor of such bias might be the following: hotel quality is unknown by consumers, but highly correlated with

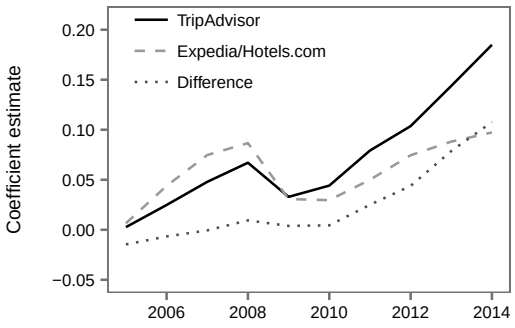


Figure 7: Coefficients on ratings by platform and ratings difference across platform.

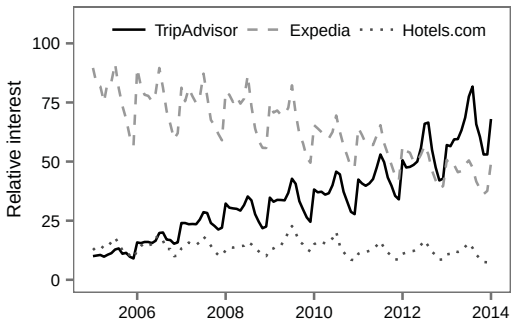


Figure 8: Google Trends by review platform.

average reviews (*i.e.*, crowdsourced reviews are generally a pretty good quality measure). Consumers learn about hotel quality through many sources: platforms such as TripAdvisor and Expedia, competing platforms not in our data (*e.g.*, hotel ratings offered by Google), their own research conducted using online and offline media, and marketing efforts, which may be disproportionately carried out by high quality hotels. Most of these information sources are omitted in our regressions, and hence appear in the error term, and yet are correlated with reviews through hotel quality.

It seems plausible to us that online review platforms are an important or perhaps even the main information source for consumers, and so although the results are biased upward, the bias may not be too large. One way to assess this is to make use of the multiple platforms in our data. To do so, we run a disaggregated version of the demand system in Equation 3, allowing the average rating on TripAdvisor $r^T_{j,t}$ to enter separately from the average rating on Expedia/Hotels.com $r^E_{j,t}$, interacting each variable with year.[5] The difference $\Delta r_{j,t} = r^T_{j,t} - r^E_{j,t}$ is on average positive in the early period of our data (*i.e.*, TripAdvisor reviews are more favorable), but is zero on average in later years. Still, there is substantial variance in ratings across platforms, with the standard deviation of $\Delta r_{j,t}$ equal to 1.75.

The estimated coefficients on $r^T_{j,t}$ and $r^E_{j,t}$ by year are plotted in Figure 7. The coefficients on TripAdvisor increase substantially over time, from a baseline of almost zero in 2005 to 0.18 by 2014. By contrast, the coefficients on Expedia display much more gradual growth, moving from 0.04 in 2006 to 0.1 by 2014. Figure 8 shows the corresponding Google Trends data for search terms related to TripAdvisor, Expedia and Hotels.com. Notice how TripAdvisor's increasing share of searches is correlated with the increase in coefficients on $r^T_{j,t}$, which is suggestive of a causal role; while the gradual decline in Expedia searches is matched by the gradual increase in the coefficients on $r^E_{j,t}$.

[5]We treat Expedia and Hotels.com as a single platform here, since they merged during our sample period (see section 4.4 below), and started showing consumers average ratings based on their joint stock of reviews.

Based on this, we run a slightly different specification in which we omit both $r^T_{j,t}$ and $r^E_{j,t}$, and instead include $\Delta r_{j,t}$ (interacted with year) as a covariate in (3). Notice that if the review platforms had no causal role — *i.e.*, average reviews are merely noisy versions of a quality measure that consumers learn from other information sources — one would expect a zero coefficient on $\Delta r_{j,t}$.[6] Instead, what we find is shown by the dotted "Difference" line in (7): a growing coefficient on the difference between the average Tripadvisor and Expedia review, negative in the first few years (when Expedia had a greater share of searches), and rising to 0.1 in the last year of our data. If one were to assume that the average Expedia review was a good proxy for all other information sources, one would conclude that an additional rating point on TripAdvisor, by itself, would increase demand in 2014 by 10%. This does suggest some causal effect of these review platforms, independent of other information sources.

4.4 Evidence on causality from a 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.[7] 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

[6] Formally, suppose $r^p_{j,t} = \gamma_{j,t} + \varepsilon^p_{j,t}$ for $p = T, E$, and that consumers know $\gamma_{j,t}$ from other sources. Then $\Delta r_{j,t} = \varepsilon^T_{j,t} - \varepsilon^E_{j,t}$, and regardless of the distribution of $\varepsilon^p_{j,t}$, there is no information for consumers in $\Delta r_{j,t}$ and hence it should play no role in determining demand.

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

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 platform, $r_{j,t}$, is by construction unchanged by the merger (the underlying set of reviews remains the same), but the average rounded rating $\tilde{r}_{j,t}$ isn't.

Using our prior notation, 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.[8] Next, define a hotel's cross-platform volume-weighted average rounded rating as:

$$\tilde{r}^P_{j,t} = \frac{\sum_{p \in P} n^p_{j,t} \tilde{r}^p_{j,t}}{\sum_{p \in P} n^p_{j,t}}. \tag{5}$$

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:

$$\Delta\tilde{r}_{j,t} = \tilde{r}^{P_2}_{j,t} - \tilde{r}^{P_1}_{j,t}, \tag{6}$$

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 equation is:

$$Q_{j,t} = \phi_1 \Delta\tilde{r}_{j,t} \times \text{Post}_t \times v_{j,t} + \phi_2 v_{j,t} + h_j + \tau_t \times m_j + \xi_{j,t}, \tag{7}$$

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 present the results on this regression in Table 3. In the first, column, we use data just from one month before and after the merger. In the third column, we use data 6 months prior to the merger to 6 months following the merger. Columns 2 and 4 include $\tilde{r}^{P_1}_{j,t}$ (the counterfactual rounded rating had the merger not happened) as an additional control. This is a robustness check in case the treatment assignment is correlated with hotel's ratings. In all four cases we estimate a statistically significant treatment effect whose magnitude, is perhaps surprisingly, equal to our prior estimates: a unit-increase in hotel's cross-platform

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

average rating is associated with 26-29% increase in demand.

In Table 4, we repeat this exercise with price as the dependent variable. Again we find a substantial effect of the ratings shock on prices. In the one-month analysis directly before and after the change, we find an effect of $12 per ratings unit, significant at 5%. The 6-month analysis finds a smaller effect of just under $7. Though this is no longer statistically significant, it is extremely close to the effect reported in Table 2 for the year 2013, which just over $7.

We have run a number of additional robustness checks on this experiment, which still need to be added to the paper. First, we have done a placebo check where we separately run the specification in (7) for each of the 6 months *prior* to the merger. In none of these placebo regressions do we find a statistically significant effect.

Second, we have checked to see that the hotels that receive a positive treatment ($\Delta\tilde{r}_{j,t} > 0$), no treatment ($\Delta\tilde{r}_{j,t} = 0$) and negative treatment ($\Delta\tilde{r}_{j,t} < 0$) are balanced on the observable characteristics (e.g. amenities, price etc). We find that they are: there are no significant differences between the three groups.

5 Counterfactuals

We would like to estimate the welfare impact of these review platforms by simulating a counterfactual world in which consumers do not have access to review data, and comparing it to the status quo as of 2014 (when review platforms appear to be most valued by consumers).

The conceptual framework was outlined earlier in Section 2. In a world without review data, consumers will form different beliefs about quality and make different choices. We already saw this in the time series in Figure 5, where we showed that consumers gradually selected hotels with higher rating (measured as of Dec 2014) over time.

In the absence of reviews, we model beliefs about quality as a linear function of observable hotel characteristics such as star rating and hotel size (but not prices, since this would introduce "price as a signal of quality" complications that we wish to avoid). Given these beliefs, we can calculate consumer choices, hotel revenue and consumer surplus (the latter according to a standard logit formula reproduced below).

But in this counterfactual world, hotels may adjust prices. Modeling this requires estimating marginal costs, and then modeling the price competition game on the supply side. We describe how we do this below.

In everything that follows, we will take as given our preferred demand estimates, reported in the second column of Table 1. Given the evidence on causality already presented, we feel reasonably comfortable in interpreting the coefficient on ratings causally. Still, as discussed

in the introduction, both the demand system and the counterfactual assume that consumers can costlessly acquire information in the current environment, and will be unable to do so in the counterfactual. This is somewhat unrealistic, and is a caveat for the work that follows.

Simulations without price responses. We assume that market size $M_{m,t}$ is equal to 1.5 times the maximum demand we observe in our data for each market-month (e.g. San Francisco in July). Based on this we can infer market shares, the log share ratio, and mean utilities $\delta_{j,t}$. Revenues under the status quo are directly observable. Consumer surplus in dollars can be calculated as

$$CS_{m,t} = \frac{1}{\alpha} \log(1 + \sum_{j \in M} \exp \delta_{j,t})$$

(up to an unidentified normalizing constant which will be differenced out when compared with the counterfactual scenario).

In the counterfactual, consumers must form beliefs based on the observable hotel characteristics alone. Write their beliefs as:

$$E[\gamma_{j,t}|x_{j,t}] = E[E[\gamma_{j,t}|r_{j,t}, x_{j,t}]|x_{j,t}] \underbrace{=}_{\text{sufficiency}} E[E[\gamma_{j,t}|r_{j,t}]|x_{j,t}] = E[\phi r_{j,t}|x_{j,t}] = \phi \Lambda x_{j,t}$$

where the reader will recall that $x_{j,t}$ is the set of observable characteristics. In the last step we assume linearity of the conditional expectation $E[r_{j,t}|x_{j,t}]$.

So to estimate the counterfactual, we first run a regression of ratings $r_{j,t}$ on $\widetilde{x}_{j,t}$ to get predicted ratings $\hat{r_{j,t}} \equiv \Lambda \widetilde{x}_{j,t}$ and then we substitute those in for the original ratings to get a counterfactual mean utility index $\delta^c_{j,t}$ for each hotel-year-month, according to:

$$\delta^c_{j,t} = x_{j,t}\beta + \alpha p_{j,t} + \phi \hat{r}_{j,t} + \xi_{j,t}$$

From this we can get counterfactual choices, revenues and consumer surplus as above.

Simulations with price responses. To allow for price responses, we need to fill in a model of the supply side. We assume that marginal costs are constant up to an idiosyncratic error:

$$c_{j,t} = \bar{c}_j + \omega_{j,t}$$

Under logit demand, the supply side first order condition can be written as:

$$s_{j,t} + \alpha s_{j,t}(1 - s_{j,t})(p_{j,t} - c_{j,t}) = 0$$

Because the error is idiosyncratic we have:

$$E[s_{j,t} + \alpha s_{j,t}(1 - s_{j,t})(p_{j,t} - \bar{c}_j)] = 0$$

and replacing the LHS of this expression with the empirical averages yields the following estimate of $\bar{c}_j$:

$$\bar{c}_j = \frac{\alpha \sum_t p_{j,t} s_{j,t}(1 - s_{j,t}) + \sum_t s_{j,t}}{\alpha \sum_t s_{j,t}(1 - s_{j,t})} \tag{8}$$

Given these costs, and the predicted ratings $\hat{r}_{j,t}$ from the previous step, we now recompute the complete information Nash equilibrium prices in each market-year-month by finding a solution to all the first order conditions given the estimated costs.

These prices are not equal on average to those we observe in the actual data, which would suggest some failure of the assumptions (e.g. demand is not logit, firms are not optimizing, firms have different information sets). Given this, we proceed by comparing the consumer surplus, revenues and prices under the *simulated* prices and status quo to the *simulated* counterfactual objects and taking differences.

Counterfactual results. See Table 5. We find that consumer surplus decreases by 66c per room-night sold in the counterfactual without a price response. This may seem like a relatively small effect, but recall that the dollarized value of an extra review point is given by ϕ_1/α, which even in 2014 is only \$15. Given our earlier evidence that such hotels have increased their prices by about \$10, the net benefit is around \$5 from choosing hotels rated one unit higher, and in fact we've shown earlier that over time consumers have only chosen hotels of about 0.1 review point higher. So the number seems reasonable, and given the number of room-nights sold each year in the United States, adds up to a substantial number (over \$1.1B).

There is little action in the revenue effects. Eliminating reviews has the effect of redistributing business from high-quality hotels to lower-quality hotels, and given that the high-quality hotels charge slightly more on average, revenue is lowered slightly by 10c a room. This hurts independent hotels more than chain/franchise hotels, in line with our earlier discussion.

Computing counterfactual price responses changes the results very little, mainly because our computed status-quo and counterfactual Nash Equilibria are very close to one another (*i.e.*, price adjustments are small). This may be an artifact of the simple substitution patterns in the logit demand system. Such effects as they are end up slightly lowering the estimated effect on consumer surplus, and slightly modifying the revenue estimates.

6 Conclusion

It is common wisdom that consumers are devoting increasing attention and time to making informed choices, accessing the many information sources at their disposal. This has implications for market structure. In this paper, we have documented the increasing correlation between online reviews and hotel bookings, and shown some evidence that popular platforms such as TripAdvisor have causal effects on purchasing behavior. We have also demonstrated that on the supply-side, hotels have responded to this phenomenon, with high-rated hotels increasing their prices, while low-rated hotels drop theirs.

This has had welfare implications. We estimate that consumers enjoy increased surplus of around $0.65 a room-night as a result of such reviews, which is a substantial amount of consumer surplus when multiplied by the number of room nights booked in the US each year. It has also increased the revenue of independent hotels relative to the chain competitors, who have disproportionately benefited from the information rich environment. Yet we have not accounted for the information acquisition costs that consumers must pay in order to read reviews and make better decisions; investigating these is an interesting direction for future research.

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.

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

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. Hortacsu. The dynamics of seller reputation: Evidence from ebay. *The Journal of Industrial Economics*, 58(1):54–78, 2010.

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

D. Elfenbein, R. Fisman, and B. McManus. Market structure, reputation, and the value of quality certification. Technical report, National Bureau of Economic Research, 2014.

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.

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.

S. Koulayev. Search for differentiated products: identification and estimation. *The RAND Journal of Economics*, 45(3):553–575, 2014.

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

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.

Table 1: Regression estimates of the impact of ratings on demand.

	(1) Log Quantity	(2) Log Quantity	(3) Log Quantity
Avg. rating	0.065***	0.006	0.059***
	(10.92)	(1.03)	(10.69)
Avg. rating × 2006		0.031***	
		(5.70)	
× 2007		0.076***	
		(9.10)	
2008		0.105***	
		(9.74)	
2009		0.043***	
		(4.51)	
2010		0.057***	
		(4.96)	
2011		0.112***	
		(7.90)	
2012		0.156***	
		(9.00)	
2013		0.202***	
		(9.55)	
2014		0.268***	
		(10.00)	
Avg. rating × Independent			0.056*
			(2.05)

Avg. rating represents the weighted cumulative average of all rating platform ratings. The dependent variable for all specifications is the log of monthly quantity rooms sold. Regressions are OLS with hotel, and market × time interaction fixed effects, and errors are clustered by hotel and market × time. Specifications are 1. Rating coefficient only 2. Rating × year interaction 3. Rating × hotel ownership interaction. Significance levels are denoted by asterisks (* p <0.1, ** p <0.05, *** p <0.01).

s

Table 2: Regression estimates of the impact of ratings on hotel prices.

	(1) Price	(2) Price	(3) Price
Avg. rating	1.481***	-0.887***	1.253***
	(6.83)	(-2.80)	(5.95)
Avg. rating × 2006		1.868***	
		(7.21)	
× 2007		3.849***	
		(9.20)	
2008		5.024***	
		(9.01)	
2009		1.192***	
		(2.78)	
2010		1.127**	
		(2.12)	
2011		2.912***	
		(4.19)	
2012		4.836***	
		(5.43)	
2013		7.153***	
		(6.56)	
2014		10.390***	
		(7.28)	
Avg. rating × Independent			2.270*
			(1.92)

Avg. rating represents the weighted cumulative average of all rating platform ratings. The dependent variable for all specifications is average monthly room price. Regressions are OLS with hotel, and market × time interaction fixed effects, and errors are clustered by hotel and market × time. Specifications are 1. Rating coefficient only 2. Rating × year interaction 3. Rating × hotel ownership interaction. Significance levels are denoted by asterisks (* p <0.1, ** p <0.05, *** p <0.01).

Table 3: Estimates of the impact of ratings on demand using the Expedia/Hotels.com merger.

Before-after window:	(1) Log Quantity (1-month)	(2) Log Quantity (1-month)	(3) Log Quantity (6-month)	(4) Log Quantity (6-month)
Rating shift	0.288**	0.292**	0.260*	0.267**
	(1.98)	(2.01)	(1.92)	(1.97)
Counterfactual rating		0.038		0.071***
		(0.59)		(2.99)

Ratings changes represent the difference between the new and old average in rounded rating displayed on the three sites after the Hotels.com/Expedia merger. Counterfactual rating is the average rounded rating had the merger never occurred. Dependent variable for all specifications is log of quantity rooms sold. Regressions are OLS with hotel, and market × time interaction fixed effects, and errors are clustered by hotel and market × time.Specifications are 1. 1 month before and after experiment 2. 1m with counterfactual control 3. 6 months before and after experiment 4. 6m with counterfactual control. Significance levels are denoted by asterisks (* p <0.1, ** p <0.05, *** p <0.01).

Table 4: Estimates of the impact of ratings on prices using the Expedia/Hotels.com merger.

Before-after window:	(1) Price (1-month)	(2) Price (1-month)	(3) Price (6-month)	(4) Price (6-month)
Rating shift	11.975**	12.170**	6.972	6.971
	(2.00)	(2.03)	(1.22)	(1.22)
Counterfactual rating		1.648		-0.016
		(0.69)		(-0.02)

Ratings changes represent the difference between the new and old average in rounded rating displayed on the three sites after the Hotels.com/Expedia merger. Counterfactual rating is the average rounded rating had the merger never occurred. Dependent variable for all specifications is average monthly room price. Regressions are OLS with hotel, and market × time interaction fixed effects, and errors are clustered by hotel and market × time.Specifications are 1. 1 month before and after experiment 2. 1m with counterfactual control 3. 6 months before and after experiment 4. 6m with counterfactual control (* p <0.1, ** p <0.05, *** p <0.01).

Table 5: Counterfactual results.

	No price response	Price response
Δ Consumer Surplus (dollars per room sold in 2014)	-0.66	-0.63
Δ Revenue (dollars per room sold in 2014)	-0.10	-0.10
Δ Revenue if chain/franchise (dollars per room sold in 2014)	-0.09	-0.09
Δ Revenue if independent (dollars per room sold in 2014)	-0.18	-0.21
Δ CS + Revenue (dollars per room sold in 2014)	-0.77	-0.74

Note: Results of the counterfactual analysis, where in the counterfactual world no ratings information is available to consumers when making choices. The units are dollars per room sold in 2014. In our data, 143 million rooms were sold in 2014. In the no price response counterfactual, prices are held fixed and market shares are recomputed; in the price response counterfactual, both market shares and prices are allowed to adjust until a Nash Equilibrium in prices is found.

A Appendix

A.1 Estimation of the price coefficient

In the main text, we note that the price coefficient in the demand system, α, is estimated flexibly using supply side moments and seasonal demand fluctuations. Here we detail exactly what this means. We assume that the residual demand curve faced by each hotel j takes the following form:

$$q_{j,t} = A_j X_{j,t} - b_j p_{j,t} + \varepsilon_{j,t}$$

for some demand shifters $X_{j,t}$. We assume the demand shifters are exogenous (i.e. $E[x_{j,t}\varepsilon_{j,t}] = 0$ for each column $x_{j,t}$ of $X_{j,t}$). We assume that the hotel-specific coefficients A_j and b_j are continuous functions of the average prices $\bar{p}_j = \frac{1}{T}\sum_{t=1}^{T} p_{j,t}$ and average quantity $\bar{q}_j \frac{1}{T}\sum_{t=1}^{T} p_{j,t}$ sold by the hotel. This allows for a very flexible specification of each residual demand curve, with the main constraint being that we require hotels that sell similar numbers of rooms at similar prices to face similar residual demands. Notice that marginal revenues are given by:

$$MR_{j,t} = A_j X_{j,t} - 2b_j p_{j,t} = q_{j,t} - b_j p_{j,t} - \varepsilon_{j,t}$$

We assume also that marginal costs take the form:

$$c_{j,t} = \bar{c}_j + \omega_{j,t}$$

where the mean-zero cost shocks $\omega_{j,t}$ are assumed iid. Now the first-order condition of the profit maximization problem requires that prices are set so that marginal revenues are equal to marginal costs. So we have that $q_{j,t} - b_j p_{j,t} - \varepsilon_{j,t} - \bar{c}_j = \omega_{j,t}$, where we have rearranged terms a little. Now suppose we have an instrument $z_{j,t}$ that is correlated with quantity $q_{j,t}$ but orthogonal to the supply shock $\omega_{j,t}$. We can use this to form the following moment condition:

$$E[(q_{j,t} - b_j p_{j,t} - \bar{c}_j)\, z_{j,t}] = 0$$

Under the assumptions that marginal costs are not seasonal, two such instruments are the expected quantity sold by hotel j in the high-season, $E[q_{j,t}|j, \text{High-season}]$, and the expected quantity sold by hotel j in the low-season $E[q_{j,t}|j, \text{Low-season}]$. We construct the empirical analogs to these objects as $\hat{q}_j^H = \frac{1}{T_H}\sum_{t\in H} q_{j,t}$ and $\hat{q}_j^L = \frac{1}{T_L}\sum_{t\in L} q_{j,t}$. Given these two instruments and a little more algebra, we can solve for b_j hotel-by-hotel as $\hat{b}_j = \frac{\hat{q}_j^H - \hat{q}_j^L}{\hat{p}_j^H - \hat{p}_j^L}$, where $\hat{p}_j^H$ and $\hat{p}_j^L$ are the average high-season and low-season prices for hotel j. But such estimates are going to be noisy, and so instead we pool, estimating $b(\bar{p}, \bar{q})$ pointwise on a

grid of $(\bar{p}, \bar{q})$ values by weighted instrumental variables regression of price on quantity, where we instrument for quantity with the interaction of hotel identity and a dummy for high season; and we weight observations using a Gaussian product kernel in $(\bar{p}, \bar{q})$-space. This delivers slope coefficients and elasticity estimates for each hotel. As shown in the paper, these are reasonably sensible, with hotels that face more competition (in terms of similarly sized hotels, or hotels that price similarly) having more elastic residual demand.

But since in the paper our demand system is a logit, with a single parameter α governing the way price affects demand, we must project down these elasticities to a lower dimensional space. Specifically, in a logit, the demand elasticity takes the form $\alpha p_{j,t}(1 - s_{j,t})$, and so we recover an estimate of α by regressing our estimated elasticities $\hat{\eta}_{j,t}$ on $p_{j,t}(1 - s_{j,t})$ without a constant, and taking the coefficient on $p_{j,t}(1 - s_{j,t})$ as our estimate of α.

A.2 Estimation of a random coefficient model

This is one way to estimate a random coefficients version of the model. We use $s^i_{j,t}$ as notation for the *inside good market shares* (i.e. rooms sold of hotel j divided by the rooms sold of all hotels in our data). We assume that in the base year, 2005, no-one used hotel ratings as the basis for decision making - and therefore made decisions based on the prices and a latent quality index $\xi_{j,t}$ that is unobserved by us. We proceed as follows:

1. Inside market shares are given in 2005 by:

$$s^i_{j,t} = \frac{\exp(\xi_{j,t} + \alpha p_{j,t})}{\sum_{k \in M_j} \exp(\xi_{k,t} + \alpha p_{k,t})}$$

where M_j is the set of hotels in the same market as j. We want to find a single ξ_j for each hotel to make the model hold as closely as possible (i.e write $\xi_{j,t} = \xi_j + u_{j,t}$ and we want to make the errors small). Fix a market m (we'll do this separately by market), and pick a numeraire hotel, denoting it as hotel 1. We normalize ξ_1 to be zero wlog. Then for any other hotel j in market m, we have:

$$\log(\frac{s^i_{j,t}}{s^i_{1,t}}) - \alpha(p_{j,t} - p_{1,t}) = \xi_j + u_{j,t} - u_{1,t}$$

and so we can get our estimates of ξ_j by computing the LHS for each hotel and taking the average (our estimate for the numerarire hotel is, once again, 0).[9]

GZ: Some hotels do not have 2005 data. What do with them? We need a ξ for them,

[9]Recall $\alpha < 0$, so the LHS will be bigger if j is more expensive than the numeraire.

but computing it based on a later entry year seems innaccuate as by then ratings do matter. Things are not too bad data-wise, because most hotels enter in 2005:

Entry year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Hotel count	3220	178	194	240	211	120	77	92	68	77

One solution is to ingore these hotels for computing w_i's but use for computing the dependent variable (inside market shares) for regression that follows in Step 6/7.

2. Post-2005, inside market shares (at a monthly level) are assumed to depend on ratings, via a random coefficient model:

$$s^i_{j,t} = \sum_n w_n \frac{\exp(\beta^n r_{j,t} + \xi_j + \alpha p_{j,t})}{\sum_{k \in M_j} \exp(\beta^n r_{j,t} + \xi_k + \alpha p_{k,t})}$$

where the weights w_n sum to one and weight a known set of random coefficients on ratings $\{\beta^n\}$. We assume that the random coeffients lie on a grid, starting at zero (i.e. $\beta^1 = 0$) and increasing by Δ each time, so that the maximum $\beta^N = \Delta(N-1)$. Notice that we're assuming that the unobservable ξ_j enters the indices in the same way as in 2005. This is an assumption, but one that seems not too bad. In this model, prices, unobservables, the price coefficient are all treated as known. We want to estimate the weights.

3. Assuming we've picked a sensible grid, now compute inside shares for each type of consumer, $s^n_{j,t} = \frac{\exp(\beta^n r_{j,t} + \xi_j + \alpha p_{j,t})}{\sum_{k \in M_j} \exp(\beta^n r_{j,t} + \xi_k + \alpha p_{k,t})}$.

4. Taking these pre-computed shares as data, we minimize a least-squares criterion function:

$$\min_w \sum_{j,t} \left(s^i_{j,t} - \sum_n w_n s^n_{j,t} \right)^2$$

This is just straight OLS, with the $(J,T) \times N$ set of shares $s^n_{j,t}$ as the independent variables and $s^i_{j,t}$ as the dependent variable.

5. We could apply this OLS estimation approach directly, but it has two problems: (a) we think the random coefficient distribution is time-varying (e.g. as people use tripadvisor more) and (b) the estimates will tend to be kind of spiky (won't look like a distribution).

6. Solution A: Estimate a two type model using only the 2014 data, where one type is uninformed (i.e. the type $\beta^1 = 0$ gets positive weight). To do this, add an additional term $\lambda \sum_{n>1} 1(w_n > 0)$ to the least-squares criterion function. This penalizes positive weight on any type other than the zero type. Dial up the hyper-parameter λ just enough

to get only two types with positive weight (the uninformed type and an informed type). Then, assuming those are the only types in the world (i.e. $\beta^1 = 0$ and $\beta^2 = \beta > 0$), go to 2013 and run constrained least squares to find the weights w_1, w_2 such that $w_1 + w_2 = 1$ that minimize the OLS criterion. Then do 2012 etc. What we should end up with is a set of year-specific weights that gradually increase the weight w_2 over time.

7. Solution B: Same as solution A (i.e estimate the model in 2014 to find which types have positive support, and then using that fixed set of types re-estimate the weights in prior years), except that we want to allow more than two types in 2014. We'd like the distribution of random coefficients to be smooth though.

   So add a penalty term $\lambda \sum_{n=1}^{N-1} \left(\log(\frac{w_{n+1}+\varepsilon}{w_n+\varepsilon}) \right)^2$ to the objective function, where ε is a small positive number there to avoid 0/0 and stuff like that. What this does is penalize sharp changes in the ratio of the weights, and so favors distributions with a connected support and those closer to uniform. Pick λ by the eyeball test - i.e. start at zero and slowly increase until the distribution of random coefficients looks reasonably smooth when plotted as a histogram. As we proceed backwards in time, we should see the weight shift towards the lower values of β.

8. Solution C: We're going to impose parametric structure, specifically a gamma distribution of coefficients. In step 2, now use a very fine grid (e.g. $\Delta = 0.01$) with a large upper bound (e.g. $\beta^N = 10$). Compute inside shares as in Step 3. Now solve a modified version of the least squares problem:

$$\min_{(\theta,k)} \sum_{j,t} \left(s^i_{j,t} - \sum_n w_n(\theta, k) s^n_{j,t} \right)^2$$

   where the weights are now constrained to be consistent with the gamma pdf:

$$w_n(\theta, k) = \frac{\frac{1}{\Gamma(k)\theta^k}(\beta_n)^{k-1} e^{-\beta_n/\theta}}{\sum_{m=1}^N \frac{1}{\Gamma(k)\theta^k}(\beta_m)^{k-1} e^{-\beta_m/\theta}} = \frac{(\beta_n)^{k-1} e^{-\beta_n/\theta}}{\sum_{m=1}^N (\beta_m)^{k-1} e^{-\beta_m/\theta}}$$

   This is now a 2D optimization problem with no regularization!

A.3 Implementation of the counterfactual

This is the procedure we followed to implement the counterfactual.

1. You can do the following steps in Stata. Make a market size assumption for each market: here M_t is the 1.5 times the maximum market size for a given market-month (month-of-year, not yearmonth).

2. Based on this, we get market shares and the share of the outside good, which in turn from the logit transformation gives us the mean utilities $\delta_{j,t} = \log \frac{s_{j,t}}{s_{0,t}}$. Define also the *gross* mean utilities $\delta^*_{j,t} = \delta_{j,t} - \alpha p_{j,t}$ (i.e. gross of price effects) — since $\alpha < 0$, these will be bigger.

3. We estimate a mean marginal cost for each hotel from (8) (just algebra). This is the cost estimate that makes the FOCs hold on average over the sample period, assuming that marginal costs are constant up to idiosyncratic error. We could do more complicated things, but this is ok.

4. We also want to estimate what consumers would have expected the ratings to be in the counterfactual world. In the counterfactual, consumers don't know ratings, but do know some observable characteristics $x_{j,t}$, and given rational expectations, the expected ratings are a linear function of those characteristics that can be found by regressing ratings on observables (when ratings are positive; i.e. omit the hotels with no review, which are coded as zero). I would use star rating and hotel age as the observables (another plausible idea is to use market fixed effects, so that consumers know the average rating of hotels in the market). The fitted values from this regression, $\widehat{r_{j,t}}$ describe the consumers counterfactual beliefs. To deal with the fact that some hotels have no reviews in the data, run a second linear regression (not a logit or probit, to avoid complications from non-linearity) to also predict the probability of having no review $\widehat{\nu_{j,t}}$ (using the same regressors).

5. Use the counterfactual ratings to compute counterfactual gross utilities, $\delta^c_{j,t} = \delta^*_{j,t} + \phi_1(\widehat{r}_{j,t} \times \widehat{\nu_{j,t}} - r_{j,t} \times \nu_{j,t}) + \phi_2(\widehat{\nu_{j,t}} - \nu_{j,t})$ (i.e. counterfactual gross utility is higher if the counterfactual rating exceeds the actual, and vice-versa).

6. At this point I would switch to Matlab. The data you need to take over are the gross utilities (both observed and counterfactual), the price coefficient, the marginal costs, market and yearmonth identiers, and chain affiliation.

7. Given the observed gross utilities and marginal costs, we want to compute a Nash Equilibrium in prices. This is a set of prices for each market and yearmonth, such that each player is best responding to all other prices. Assuming no joint ownership (i.e. in each market, each hotel is owned by a different player), and dropping the t

notation, this is a price vector $\mathbf{p}$ of length N_m, with the property that $s_j(\mathbf{p})+\alpha s_j(\mathbf{p})(1-s_j(\mathbf{p}))(p_j-c_j)=0$ for all j, where c_j is the marginal cost estimate for hotel j. Since s_j depends on $\mathbf{p}$ via $s_j=\frac{\exp(\delta_j^*+\alpha p_j)}{1+\sum_k \exp(\delta_k^*+\alpha p_k)}$, this is a non-linear set of N_m equations that must be solved via some solver. Do this to return an equilibrium price vector $\mathbf{p}^*$.

8. If we want to be more careful, we should correct for joint ownership. We will solve N_m equations as before, but the FOC is now:

$$s_j(\mathbf{p})+\alpha s_j(\mathbf{p})(1-s_j(\mathbf{p}))(p_j-c_j)-\alpha s_j(\mathbf{p})\sum_{k\in\mathcal{J},k\neq j} s_k(\mathbf{p})(p_k-c_k)=0$$

   where the notation in the last summation means the sum over all other hotels with the same owner as hotel j.

9. Repeat the same exercise with counterfactual gross utilities δ_j^c to get a counterfactual equilibrium price vector $\mathbf{p}^c$.

10. At this point, we now have gross utilities and equilibrium prices under the status quo and the counterfactual. We want to compute for each market the consumer surplus and hotel profits (or maybe profits of independents and chains separately).

11. Get the simulated status-quo market shares as $s_j=\frac{\exp(\delta_j^*+\alpha p_j^*)}{1+\sum_k \exp(\delta_k^*+\alpha p_k^*)}$. Ditto for counterfactuals, using δ_j^c and p_j^c. Get per hotel profits as price less marginal costs times market share times market size $M_{m,t}$, add them up to desired level - profits of independents in San Francisco in 2014, e.g.

12. To get consumer surplus, we need to simulate some consumers. Consumers can choose between any of the N_m hotels in a market and the outside option of non-purchase. The outside option gives mean utility of zero; the others give $\delta_j^*+\alpha p_j^*$ (in the status quo), and $\delta_j^c+\alpha p_j^c$ (in the counterfactual). Draw S consumers by sampling a $(N_m+1)\times S$ matrix of random variables from a standard Gumbel distribution (see wikipedia). For each consumer, pick the good that gives the highest utility, where utility is the sum of mean utility and idiosyncratic shock. This is itself a $N_m+1\times S$ matrix with a single positive entry in each column. Compute total consumer surplus as the realized utility for each of the S agents; divide by S to get the average consumer surplus. Multiply by market size to get things on the same scale as the hotel profits. We should have for each market a status quo and counterfactual consumer surplus.

13. All of the above analysis assumed our dataset contained all the hotels in the market. But before we get to Matlab, we could try impute gross utilities (observed and unob-

served) and marginal costs for the census data. The easiest way to do this is some sort of matching procedure: for each hotel in the census but not the main data, condition on some observables (independent/chain, star rating), get all the hotels in the main dataset that match on those observables, and sample uniformly at random from the matched set to get imputed values of (gross utility, marginal cost) for that hotel. Repeat for each hotel in the census but not main dataset to "complete" the data. Then proceed as before. This is pretty straightforward, and easy to explain. Does it strike you as sensible?