

# Analyzing Signaling and Reputational Effects of Customer Online Reviews on Hotel Revenues

## Group Project Proposal - Survivals

### 1. Member & Member Roles

Member	Contribution
Shuyi Zhang	Web Scraping TripAdvisor Data
Joyce Fu	Web Scraping Booking.com Data
Sitong Zhang	Data Cleaning & Data Processing
Guanhong Liu	Data Analysis

All team members share the same responsibility for the final report and video.

### 2. Project Background

The technological advancement in the twenty-first century has led to the digitization of interactions between firms and consumers on online platforms, consumers nowadays can rate or write reviews on platforms to express their customer experiences. These ratings and reviews are the “signals” of product quality. Accumulation of positive ratings and reviews also help build reputation, increasing sales and a company’s revenue.

Online reviews have been considered more effective in influencing consumer choice than traditional marketing and mass advertising (Gretzel and Yoo, 2008; Yang and Mai, 2010; Ye et al., 2011; Zhang et al., 2010). In this case, it has been argued that there is a positive effect of customer ratings on firm performance (Luca, 2016; Yang et al., 2018). However, little is known about the empirical mechanisms whereby firm performance is affected by customer ratings.

To explore the effect of customer ratings on firm performance, this project builds on reputation theories. In reputation theories (Klein and Leffler, 1981; Shapiro, 1982), there are two effects of reviews/ratings on firm performance: 1) reputational effect and 2) signaling effect. Reputation

effect is one company is incentivized to build reputation to set higher prices or increase customer loyalty and purchases to gain more profits. Signaling effect is the reviews/ratings can implicate product quality, reducing the probability of adverse selection. We will collect hotel revenue and hotel rating and reviews data for the model estimation and develop empirical methods to identify and estimate the practical reputational effects and signaling effects, lending empirical support to reputation theories.

### **3. Project goals**

This project aims to investigate (1) reputational effects and (2) signaling effects of customer reviews on hotel rentals websites(TripAdvisor and Booking.com) on Texas hotel revenues.

### **4. Research Questions**

- a. Do customer ratings on TripAdvisor affect Texas hotel revenues?
- b. If so, how do the customer ratings on TripAdvisor affect Texas hotel revenues? Through signal or reputation channels?
- c. How great are the reputational effects and signaling effects?
- d. Do the effects also exist in other online housing rental markets(Booking.com)?  
(Robustness Check)

### **5. Steps**

- a. Collecting Data

We plan to collect data from 3 datasets:

- Hotel Tax Receipts:

This monthly hotel occupancy tax data is a public record provided by the Texas government. The dataset can be directly downloaded from the Texas Comptroller of Public accounts

(<https://data-secure.comptroller.texas.gov/main/files/public-files>). In this monthly data, it includes taxpayer name, address, location and total room receipts(i.e. hotel revenues). The variables we are interested in are monthly room revenues of hotels in Texas.

- The second data source is TripAdvisor.com (<https://www.tripadvisor.com/>), a pioneer of online travel intermediation specializing in hotel booking. Typically, a verified user can post reviews about his or her experience with various

accommodation products, ranging from hotel, motel, bed and breakfast to other vacation rentals. Also, anyone, with or without an account with TripAdvisor, can access reviews posted by other consumers and obtain a wide range of other publicly available information about hotel characteristics that are listed on TripAdvisor.com. There are two ways for users to comment on their experience: 1) rating a hotel or a restaurant on a scale of 1 star to 5 stars, or/and 2) posting a textual review about their experience. Our group would perform web scraping to get this dataset. The information we are interested in include: hotel names, addresses, amenities (such as free parking, free breakfast, etc.), online review data (including ratings and comments).

- The third dataset is similar to the second one. We would scrap hotel and customer data from Booking.com. Similar information as the Tripadvisor dataset will be included in this dataset.

#### b. Data Cleaning & Data Processing

Firstly, for data cleaning, we would remove all the missing values and irrelevant values. For instance, in the first dataset, besides TX, GA is also included in the Taxpayer State field. We would remove all GA data. Another example is that for some hotels, instead of monthly data, only quarterly data is included in the dataset. For those data points, we plan to evenly divide the quarterly data into month data.

To combine data from TripAdvisor, Booking.com and hotel revenue data from the Texas Comptroller of Public Accounts, we plan to use hotel addresses to match the same hotels.

#### c. Data Analysis

For data analysis, we first plan to present a preliminary description of the relationship between monthly hotel revenues and customer rating from TripAdvisor. We hypothesize that hotel revenue should be positively associated with Tripadvisor's customer rating on a certain interval.

Then, our group plans to proceed to build regression models and explain this positive relationship, and thereby examine the signaling effects and reputational effects. To be specific, we would use a fixed effects regression to estimate the relationship between

hotel revenues and TripAdvisor's customer ratings and a regression discontinuity specification to estimate the causal effect. We will use hotel revenue as the dependent variable and use independent variables including overall ratings presented on websites(reputation effect), the actual average ratings(signal effect) and other control factors.

Thirdly, we proceed to conduct a robustness check by performing the same analysis steps on Booking.com customer rating data & hotel revenue data from the Texas Comptroller of Public Accounts. One of our concerns is that the customer rating data can be manipulated, for example, hotel staff can post fake reviews to inflate their hotel ratings. We plan to compare the analysis results from (Tripadvisor + monthly hotel revenues) and (Booking.com + monthly hotel revenues). If the results are similar, we can then prove our reputational and signaling effect.

## 6. Reference

- a. Gretzel, U., Yoo, K.H., 2008. Use and impact of online travel reviews. In: O'Connor, P., Hopken, W., Gretzel, U. (Eds.), *Information and Communication Technologies in Tourism 2008*. Springer, Vienna, pp. 35–46.
- b. Klein, B., Leffler, K.B., 1981. The role of market forces in assuring contractual performance. *J. Political Econ.* 89 (4), 615–641.
- c. Luca, M., Zervas, G., 2016. Fake it till you make it: reputation, competition, and Yelp review fraud. *Manag. Sci.* 62 (12), 3412–3427.
- d. Shapiro, C., 1982. Consumer information, product quality, and seller reputation. *Bell J. Econ.* 13 (1), 20–35.
- e. Yang, J., Mai, E., 2010. Experiential goods with network externalities effects: an empirical study of online rating systems. *J. Bus. Res.* 63 (9–10), 1050–1057.
- f. Yang, Y., Park, S., Hu, X., 2018. Electronic word of mouth and hotel performance: a meta-analysis. *Tour. Manag.* 67, 248–260.
- g. Zhang, J.Q., Craciun, G., Shin, D., 2010. When does electronic word-of-mouth matter? A study of consumer product reviews. *J. Bus. Res.* 63 (12), 1336–1341.