Nakamura, Akihiro

Conference Paper

Evaluating Customer Reviews in Matching Services on the Internet

28th European Regional Conference of the International Telecommunications Society (ITS): "Competition and Regulation in the Information Age", Passau, Germany, 30th July - 2nd August, 2017

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Nakamura, Akihiro (2017) : Evaluating Customer Reviews in Matching Services on the Internet, 28th European Regional Conference of the International Telecommunications Society (ITS): "Competition and Regulation in the Information Age", Passau, Germany, 30th July - 2nd August, 2017, International Telecommunications Society (ITS), Calgary

This Version is available at:
https://hdl.handle.net/10419/169486

Evaluating Customer Reviews in Matching Services on the Internet

Akihiro Nakamura
Yokohama City University

We can now buy almost anything via the Internet. Prevailing matching services also let us buy services via the Internet. We can book hotels, restaurants, or taxi services by visiting portal sites and searching for a suitable one among various alternatives. Consumers use customer review information when buying something via the Internet. Of course, she/he might also use this information when purchasing something in an actual store. Therefore, most shopkeepers worry about their customer reviews; otherwise, their businesses could fail. This is also true for service providers in charge of internet businesses. For example, UBER drivers cannot continue as drivers if their customer reviews are poor. Thus, customer reviews are increasingly important to internet businesses.

Customers search for the best service providers through internet portal sites, particularly for restaurants, hotels, and taxi services, which may be being provided via the Internet for the first time. In Japan, with the exception of restaurants, hotel and taxi services have begun to discuss changing their portal site regulations after the emergence of internet services such as Airbnb and UBER. These services are a type of matching service connecting customers and providers, who provide their services as non-professionals. Thus far, the professional hotel and taxi markets have been regulated for safety reasons in Japan. However, these regulations are too strict to be fulfilled by “individual” service providers.

For more efficient resource usage in a shared economy, the government started to consider a new regulatory framework for these service providers to establish a new safety standard. The platformers providing matching services, such as UBER and Airbnb, emphasize the importance of customer reviews, which play an important role for the emergence of a competitive market where only good service providers survive. However, currently there is not so many researches relating customer reviews and competition. This study tackles this topic.

In this paper, we assessed the monetary value of customer reviews in three Internet service industries: restaurants, hotels, and taxis. We conducted an internet survey in March 2017 that measured the WTP for choosing a good/average/bad-rating service provider as compared to a no-rating service provider for restaurants, hotels, and taxi services. In addition to measuring WTP for these services, by employing conjoint analysis frameworks, we analyze whether the consumer reviews work as replacement of the current taxi service regulations, such as management of drivers, taxi drivers licenses, and so on.

The results reveal the following: First, consumers choose a service provider with a “good” rating even if its price is 21% higher for restaurants, 9% higher for hotels, and 5% higher for taxi

services than that of a provider with average ratings. Second, consumers are willing to choose a service provider with a "bad" rating if its price is 20%, 18%, and 14% lower for restaurant, hotels, and taxi services, respectively, than that of a provider with average ratings. Third, consumers choose a service provider that is not rated if its price is 24%, 23%, and 17% lower than that of a provider with average ratings for restaurants, hotels, and taxi services, respectively. Fourth, our conjoint analysis results show that it's difficult to replace the government regulations in Japanese taxi market by the consumer reviews.

2. Stated Preference Experiments

This section briefly explains the data used in this study. In March 2017, a survey was conducted to collect information about consumers' service choice behavior by knowing customers' reviews from a hypothetical internet matching site. Consumers' service choice behavior can be examined by analyzing data such as revealed preference (RP); however, this study used a stated preference (SP) survey to capture consumer preferences. One of the reasons why we employ an SP survey is data availability. Although there are few studies that deal with consumer reviews, such as Proserpio and Zervas (2016), it is still difficult to evaluate consumer review information using real data.

This paper employs two approaches to evaluate the monetary values of consumer reviews. The data collection methods are explained in the following sections.

(1) Data Collection for Estimating WTP for Choosing a Different Rating Service

The first approach is to measure the willingness to pay (WTP) for a good/average/bad-rated service provider compared with a non-rated service provider of restaurants, hotels, and taxi services. To collect data for estimating WTP, sequential questions were posed asking how much the respondent would pay for each hypothetical alternative compared with the status quo alternative, rated as "average."

More concretely, we asked our respondents to answer the following questions:

Q1: Imagine the situation that you plan to have dinner with your friend and try to book a table at a restaurant where you and your friend have never been, through internet portal sites.

You find the following two alternative restaurants:
1. Restaurant A, which is rated 4 points. The average rating score is 4 points. The cost (for one person) of this restaurant is JPY5,000.
2. Restaurant B, which is rated 4.2 points. The average rating score is 4 points.

Quote the cost (for one person) for Restaurant B that would permit you to choose Restaurant B (good-rated restaurant).

Note: the above two restaurants have almost the same location, atmosphere, and foods but not cost (for one person). Each provides the information on their internet homepage.
You and your friend have never visited either of the two restaurants.

1) I will choose Restaurant B only if the cost (for one person) of Restaurant B is the same as Restaurant A (JPY5,000).
2) I will choose Restaurant B if the cost (for one person) of Restaurant B is under JPY5,250.
3) I will choose Restaurant B if the cost (for one person) of Restaurant B is under JPY5,500.
4) I will choose Restaurant B if the cost (for one person) of Restaurant B is under JPY5,750.

…

11) I will choose Restaurant B if the cost (for one person) of Restaurant B is under JPY7,500.
12) I will choose Restaurant B even if the cost (for one person) of Restaurant B is above JPY7,501. Fill in your concrete maximum cost (for one person) for Restaurant B:_____.

Q2):

This question relates to a similar situation as above. A new (third) alternative restaurant appears in your search, Restaurant C, rated 3.8 points.

Quote the cost (for one person) for Restaurant C that would permit you to choose Restaurant C (bad-rated restaurant).

1) I am OK with Restaurant C even if the cost (for one person) of Restaurant C is the same as Restaurant A (JPY5,000).
2) I will choose Restaurant C if the cost (for one person) of Restaurant C is under JPY4,750.
3) I will choose Restaurant C if the cost (for one person) of Restaurant C is under JPY4,500.
4) I will choose Restaurant C if the cost (for one person) of Restaurant C is under JPY4,250.

…

11) I will choose Restaurant C if the cost (for one person) of Restaurant C is under JPY2,500.
12) I will choose Restaurant C if the cost (for one person) of Restaurant C is below JPY2,499. Fill in your concrete maximum cost (for one person) for Restaurant C:_____.

Q3):

This question relates to a similar situation as above. A new (fourth) alternative restaurant again appears in your search results, Restaurant D, which is non-rated.
Quote the cost (for one person) for Restaurant D that would permit you to choose Restaurant D (non-rated restaurant).

1) I am OK with Restaurant D even if the cost (for one person) of Restaurant C is the same as Restaurant A (JPY5,000).
2) I will choose Restaurant D if the cost (for one person) of Restaurant C is under JPY4,750.
3) I will choose Restaurant D if the cost (for one person) of Restaurant D is under JPY4,500.
4) I will choose Restaurant D if the cost (for one person) of Restaurant D is under JPY4,250.

…

11) I will choose Restaurant D if the cost (for one person) of Restaurant D is under JPY2,500.

12) I will choose Restaurant D if the cost (for one person) of Restaurant D is below JPY2,499. Fill in your concrete maximum cost (for one person) for Restaurant D:_____.

The above three questions capture each respondent's WTP for differently rated services, such as a restaurant, hotel, and taxi. As baseline figures, we set the service prices of average-rated restaurants, hotels, and taxis as JPY5,000, 10,000, and 2,500, respectively. The respondents compare this price with providers rated differently and state their WTP, which allows us to calculate WTP for each provider.

(2) Conjoint Analysis Approach

The second approach is to conduct a conjoint analysis focusing on the taxi service market. The purpose for this analysis is to discuss whether customer reviews can be used as alternative tools for regulating the Japanese taxi market. Currently in Japan, there are various regulations regarding taxi services and drivers need a taxi driver's license, which assures more advanced and safe driving skills. In addition to regulations on drivers, there are regulations on taxi companies regarding drivers' working hours and safety facilities. This is due to the predominance of company-based taxi services in Japan, with few independent operators. These regulations deal with safety aspects for providing taxi services. Since consumer reviews reflect various service qualities, consumers might notice not only service quality but also driving technique or safety aspects when they travel by taxi. Consumer review scores might reflect driving technique/safety regardless of whether the driver has a taxi license or not. If so, then consumers might choose the taxi even if its driver did not have a taxi license based on very good consumer review. In this analysis, we analyze how consumer review

scores can substitute each regulation.

We employ SP approaches to collect data for this analysis. This is because, in Japan, taxi services are not provided without regulations. For instance, UBER services without a taxi driver's license are not permitted.

By designing a choice experiment, researchers could assure the variability of attribute levels, including price, and avoid collinearity among attributes. These are the advantages that SP data offer over RP data. This study's SP survey used a conjoint questionnaire. Conjoint analysis as an SP experimental technique has been applied in an array of disciplines. Hensher (2001, 2004) applied it to automobile travel. Layton (2000) conducted environmental research and Kim (2005), Lee et al. (2006), and Nakamura (2010a, 2010b) analyzed mobile phone demand using conjoint analysis. Marketing research frequently uses conjoint analysis (Huber & Train, 2001), in which researchers construct hypothetical bundles of attributes that describe a product or service and ask respondents to state their preferences from the hypothetical alternatives.

This study's conjoint questionnaire comprises attributes related to Japanese taxi market regulations. Shinkeiren (2016), who promoted ride sharing services in Japan in their proposal, stated that these regulations can be replaced by other tools, including competitive pressure through customer reviews.

The range of attributes and levels in each alternative in the experiment appear in Table 1. Figure 1 illustrates an example of the conjoint questionnaire.

Table 1: Design of Conjoint Analysis

	Levels in the Current Card	***Levels of Two Alternatives***
Consumers' reviews	No rating	No rating, 3.8 points, 4.0 points, or 4.2 points
Drivers' working-hour management	Managed	Managed or Unmanaged
Criminal/Accidental records check	Checked	Checked or Unchecked
Taxi driver's license	Hold	Hold or Not hold
Fare (JPY)	2,500	1,500, 2,000, 2,500, 3,000, or 3,500

In this analysis, each alternative is bundled according to five attributes: (1) consumers' review scores including no rating, (2) whether companies/portal sites manage drivers' working hours, (3) whether companies/portal sites check/manage criminal/accidental background, (4) whether a driver has a taxi driver's license or not, and (5) the fare for each taxi service. Each experiment listed three alternatives: a card that reflects the current Japanese taxi market's regulations (no consumer review and all regulations exist) and two alternative cards where some regulations "unexist" with

various consumers' reviews (Figure 1).

We bundle one status quo alternative that reflects current Japanese taxi services and two hypothetical alternative cards and asked the respondents to order them.

	Current Card (Stats Quo Card)	***Alternative 1***	***Alternative 2***
Customer Reviews	no rating	no rating	3.8 points
Management of Driver's Working-hour	Managed	Unmanaged	Managed
Criminal/Accidental Records Check	Checked	Checked	Unchecked
Taxi Driver's License	Taxi drivers license	No (only general drivers license)	Taxi drivers license
Fare (JPY)	JPY 2,500	JPY 1,500	JPY 1,500
Most Prefereble	□	□	□
Most Unprefereble	□	□	□

Figure 1: Example of conjoint analysis questionnaire

Participants were taken from a survey panel organized by an internet survey company (Macromill Research, Inc.). They responded to the abovementioned five rank-ordered choice questions. Each experiment listed the aforementioned three alternatives. A sample of 1,450 responses was obtained. Table 2 presents the basic statistics of the dataset.

The number of profiles would have become unwieldy, had all possible combinations of attributes been considered. Therefore, the survey's conjoint profiles were narrowed to 40 patterns using orthogonal design methods, considering each main effect and possible interactive effects (see Louviere et al., 2000; Kuehl, 1999; Hensher et al., 2005).

Table 2: Basic Statistics

Age		Gender		Household Income/year		Individual Income/year	
20-24	83	Male	724	Under JPY 2millions	113	Under JPY 2millions	521
25-29	142	Female	726	2-4Millions	291	2-4Millions	324
30-34	136			4-6Millions	337	4-6Millions	185
35-39	144			6-8Millions	180	6-8Millions	88
40-44	171			8-10Millions	113	8-10Millions	59
45-49	176			10-12Millions	56	10-12Millions	16
50-54	136			12-15Millions	34	12-15Millions	7
55-59	132			15-20Millions	23	15-20Millions	8
over60	330			over20Millions	9	over20Millions	3
				No answer	294	No answer	239
total	1450	total	1450	total	1450	total	1450

3. Econometric Approach

For analyzing the conjoint data, this study's consumer behavior model is based on the

random utility framework proposed by McFadden (1974). Assuming that customer *i* faces a choice among *J* alternatives in each of the *T* choice sets, the utility functional form when individual *i* chooses alternative *j* in choice set *t*, where x_{jt} is expressed as a vector of independent variables, is

$$U_{ijt} = \beta' x_{jt} + \varepsilon_{ijt} .$$

This study assumed three alternatives. The distribution of random disturbance ε_{ijt} is assumed to be an independent and identical extreme value: the unknown coefficient vector β.

More concretely, the assumed utility function is

$$\begin{aligned} U_{ijt} = {} & \beta_{NR} NORATED_{jt} + \beta_{CR} CONREV_{jt} + \beta_{MWH} MWH_{jt} \\ & + \beta_{CAC} CRACC_{jt} + \beta_{TL} TAXLIC_{jt} + \beta_{FR} FARE_{jt} + \varepsilon_{ijt} \end{aligned} .$$

The independent variables correspond to the attributes in the experiment (Table 1 and Figure 1). *NORATED* is a dummy variable that takes the value 1 if a taxi has no rating and 0 otherwise. *CONREV* stands for consumer review score. *CONREV* takes the value 4.0 (the average rating score) if a taxi has no rating; therefore, the coefficient of *NORATED* stands for the negative effect of no rating from the average-rated taxi service. *MWH* takes the value 1 if a taxi driver is managed in his/her working hours and 0 otherwise. *CRACC* takes the value 1 if a taxi driver is checked about his/her criminal or serious accident background and 0 otherwise. *TAXLIC* takes the value 1 if a driver has a taxi driver's license and 0 otherwise. *FARE* stands for taxi fare in JPY.

Our survey data are contingent ranking conjoint data. The model makes full use of all ranking information by repeatedly applying the conditional logit model. Each choice set consists of a first-ranked choice and lower-ranked alternatives.[1] The probability of individual *i*'s observed sequence of rankings is expressed as

$$L\left(r_i = \{r_{i1}, \cdots, r_{iT}\} \middle| \beta\right) = \prod_{t=1}^{T} \prod_{m=1}^{J-1} \frac{e^{\beta' x(r_{mt})}}{\sum_{k=m}^{J} e^{\beta' x(r_{kt})}},$$

where r_{it} is the vector of individual *i*'s ranking responses of choice set *t* and $x(r_{mt})$ is the vector of independent variables of the alternative ranked *m* in descending preference; that is, we employ the rank-ordered conditional logit model.

[1] Hausman and Rudd (1987) pointed out the possibility that a respondent in a survey will pay more careful attention to her top choice or top few choices rather than carefully ranking all alternatives. In this sense, there exists the tradeoff that using more ranks gives more efficient parameter estimates, but it can also introduce a bias in the results. Chapman and Staelin (1982) suggest to only use the first few ranks in the estimation. Hausman and Rudd develop the statistical test for this type bias and also propose alternative estimation methods. However, because there are only three alternatives including the status quo alternative in our survey, we assumed respondents pay same attentions to the twice rankings.

4. Estimation Results

Before estimating the rank-ordered conditional logit model for the possibility of substitutability between consumer reviews and government regulations, we first estimate the difference in WTP for the differently rated services using data as explained in the earlier section.

Table 3: Estimation Results of WTP

		Provider with No Rating	*Provider with a "Bad" Rating*	*Provider with Average Ratings*	*Provider with a "Good" Rating*
WTP (Based on the WTP for one with average ratings)	***Restaurant***	JPY 3,785	JPY 4,012	JPY 5,000	JPY 6,048
	Hotel	JPY 7,700	JPY 8,225	JPY 10,000	JPY 10,864
	Taxi	JPY 2,079	JPY 2,155	JPY 2,500	JPY 2,614
Ratio of difference in WTP from the one with average ratings	***Restaurant***	-24.3%	-19.8%	100.0%	21.0%
	Hotel	-23.0%	-17.7%	100.0%	8.6%
	Taxi	-16.9%	-13.8%	100.0%	4.6%

Table 3 shows the estimated average WTP for each differently rated service. Since the survey directly asked for respondents' WTP for each differently rated service, the figures in the upper rows of Table 3 are calculated to take the average of all respondents' stated WTP. The shaded column in Table 3 contains the baseline figures of average-rated services. We set the service prices of average-rated restaurants, hotels, and taxis as JPY5,000, 10,000, and 2,500, respectively. The lower rows show the percentage differences from WTPs for each average-rated service.

Based on Table 3, the effects of difference in rating vary among services. As for restaurants, the difference-in-difference between "bad" and "good" is almost symmetrical; however, the other two services show asymmetrical effects on WTP for "bad" and "good" services. For example, consumers would pay 8.6% more for a "good" hotel over an "average" hotel, but only 82.3% of an "average" hotel price for a bad-rated hotel. In addition, they would pay 4.6% more for a "good" taxi than an "average" taxi, but only 84.2% of an "average" taxi fare for a bad-rated taxi. As non-rated services, taxi services are different from other services. Currently, customer reviews of hotels and restaurants are common on internet portal sites; however, this is not the case with taxi services. Consumers also might feel that non-rated hotels and restaurants have some problems. A common feature is that non-rated providers of these three services attract the lowest WTP among all providers.

Next, we move to the estimation results of conjoint analysis for replicability of consumer reviews with regulations. Table 4 presents the estimation results and significant estimates with appropriate signs. Each coefficient of the current regulations' dummy variables, such as *MWH*, *CRACC*, and *TAXLIC*, takes a significantly positive value. The coefficient of consumers' review scores exhibits a significantly positive value. The coefficients assumed as negatively evaluated, such as those of *FARE* and *NORATED*, take significantly negative values.

Table 4: Estimation Results

Rank-Ordered Logit	Coef.	p-value	WTP
No Rating Provider Dummy Variable (*NORATED*)	-0.121	(0.000)	JPY - 405
Customer Review Score (*CONREV*)	0.252	(0.000)	JPY 841
Management of Driver's Working-hour (*MWH*)	0.214	(0.000)	JPY 714
Criminal/Accidental Background Check (*CRACC*)	0.286	(0.000)	JPY 955
Taxi Driver License (*TAXLIC*)	0.262	(0.000)	JPY 876
FARE	-0.299	(0.000)	

Next, this study considers the relative importance of each attribute change. The relative importance between two attribute changes is calculated as the ratio of the attributes' coefficients. The coefficient ratio for each attribute to the coefficient of the monetary attribute is interpreted as WTP in compensating variation. *FARE* is used in this estimation model. Therefore, WTPs based on this variable can be calculated.

The right-hand column in Table 4 reports the calculated WTP for each attribute change. The results reveal that the value of a one-point difference in consumer reviews is JPY841, which is almost the same as each existing regulation, such as the management of working hours, criminal/serious accidental background check, and holding a taxi driver's license. Non-rated providers can be undervalued by JPY405, which is almost the same as the result of WTP in the above direct calculation. However, we need to take into account the meaning of a one-point difference in consumer review scores, which is a very large difference. In this sense, consumer review scores cannot replace government regulations because we rarely find a one-point customer review difference in portal sites of other service matching sites.

6. Concluding Remarks

In this paper, we assessed the monetary value of customer reviews in three internet service industries: restaurants, hotels, and taxis. We conducted an internet survey in March 2017 that measured the WTP for good/average/bad-rated service providers compared with non-rated restaurants, hotels, and taxi services. In addition to measuring WTP for these services, by employing conjoint analysis frameworks, we analyzed whether consumer reviews could replace current taxi service regulations, such as driver management and taxi licenses.

The results reveal the following. First, consumers choose a service provider with a "good" rating even if its price is 21% higher for restaurants, 9% higher for hotels, and 5% higher for taxi services than that of a provider with average ratings. Second, consumers are willing to choose a service provider with a "bad" rating if its price is 20%, 18%, and 14% lower for restaurants, hotels, and taxi services, respectively, than that of a provider with average ratings. Third, consumers choose a service provider that is not rated if its price is 24%, 23%, and 17% lower than that of a provider

with average ratings for restaurants, hotels, and taxi services, respectively. Fourth, our conjoint analysis results show that it is difficult to replace government regulations in the Japanese taxi market with consumer reviews.

Acknowledgments

This research was partly supported by a Grant-in-Aid (No. 2430089) from the Japan Society for the Promotion of Science.

References

Chapman, R.G., Staelin, R. 1982. Exploiting rank ordered choice set data within the stochastic utility model, Journal of Marketing Research 19, pp.288-301.

Hausman, J., Ruud, P., 1987. Specifying and testing econometric models for rank-ordered sata. Journal of Econometrics 34, pp.83-104.

Hensher, D.A., 2001. Measurement of the valuation of travel time savings. Journal of Transport Economics and Policy 35, pp.71–98.

Hensher, D.A., 2004. Identifying the influence of stated choice design dimensionality on willingness to pay for travel time savings. Journal of Transport Economics and Policy 38, pp.425–446.

Huber, J., Train, K., 2001. On the similarity of the classical Bayesian estimates of individual mean partworths. Marketing Letters 12, pp.257–267.

Kim, Y., 2005. Estimation of consumers' preferences on new telecommunications services: IMT-2000 service in Korea. Information Economics and Policy 17, pp.73–84.

Layton, D.F., 2000. Random coefficient models for stated preference surveys. Journal of Environmental Economics and Management 40, pp.21–36.

Lee, J., Kim, Y., Lee, J.D., Park, Y., 2006. Estimating the extent of potential competition in the Korean mobile telecommunications market: Switching costs and number portability. International Journal of Industrial Organization 24, pp.107–124.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Eds), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.

Nakamura, A., 2010a. Changes in consumers' behavior when a vertically integrated service is separated -The case of Japanese mobile phone services-, Economics Bulletin 30, pp. 437-449.

Nakamura, A., 2010b. Estimating switching costs involved in changing mobile phone carriers in Japan: Evaluation of lock-in factors related to Japan's SIM card locks. Telecommunications Policy 34, pp. 736-746.

Proserpio, D, Zervas, G., 2016. Online reputation management: Estimating the impact of management responses on consumer reviews, Marketing Science, Forthcoming.

Shinkeiren, 2016, For the ride sharing. (in Japanese).