경영학 석사학위논문

Rating Environments and Consumers' Propensity to Engage in Ratings

리뷰 작성 환경과 소비자들의 리뷰 작성 경향

2019년 8월

서울대학교 대학원

경영학과 마케팅전공

김 은 선

Rating Environments and Consumers’ Propensity to Engage in Ratings

지도교수 송 인 성

이 논문을 경영학 석사학위논문으로 제출함

2019년 7월

서울대학교 대학원

경영학과 마케팅 전공

김 은 선

김은선의 경영학 석사학위논문을 인준함

2019년 7월

위 원 장	주 우 진 (인)
부위원장	박 기 완 (인)
위 원	송 인 성 (인)

ABSTRACT

Do previous rating environments affect consumers' propensity to engage in subsequent ratings? Whereas prior research has addressed the relationship between post purchase evaluation and the incidence decision of expressing opinions, little work has examined the underpinnings of the link between posted reviews and subsequent incidence decisions. Using a large dataset of restaurant reviews collected from Yelp.com, I investigate social dynamics in the opinion expression. The objective of this research is to examine the systematic link between prior and posterior reviews and reveal the factors that are associated with the aggregate number of reviews in the subsequent period. The factors affecting the consumers' incidence decisions have the potential to systematically alter the compositions of opinions in the review websites. This paper examines the self-selection in the consumers' decision to contribute to the online conversation by empirically identifying systematic biases in review websites by studying restaurant reviews at the content level. I present the following findings of the relationship between previous rating environments and subsequent review generation: (1) more reviews are contributed toward the restaurants with more reviews in the previous period, (2) activists contribute more reviews toward the restaurants with the fewer cumulative number of reviews in the previous period, (3) more reviews are contributed toward the restaurants with higher Yelp rating in the previous period, and (4) more reviews are contributed toward the restaurants with a shorter average length of reviews in the previous period. Overall, these results show that online reviews are disproportionately written for the specific rating environments with a consistent pattern of reviewers responding to previously posted reviews.

Keywords: online word of mouth, social dynamics, social influence, rating environments, propensity to engage in ratings
Student number: 2017-24741

Contents

1. INTRODUCTION

More and more consumers have voluntarily voiced online in social media environments where consumers can broadcast opinions to a broad audience. Currently, numerous review websites are easily accessible to acquire information about the purchase experience from a multitude of other consumers. This paper provides empirical evidence on the impact of online rating environments on the consumers' decision to participate in the online conversation.

According to a 2016 Pew Research Center report, 82 percent of U.S. adults say that they read online customer ratings or reviews before purchasing items for the first time. However, only about 43% post their reviews about products they have bought, the restaurants they have visited and the services they have used.[1] Most people do not write a review; they read them.[2]

The low frequency of the consumers' contribution has prompted considerable research on the motivations of users to express opinions on

[1] For more specific results on a survey on online reviews, see https://www.pewinternet.org/2016/12/19/online-reviews.

[2] Yelp Blog (June 7), https://blog.yelp.com/2011/06/yelp-and-the-1-9-90-rule (accessed June 9, 2019).

review websites. Schlosser (2005) examined the behavioral difference between posters (those communicating their experience to others) and lurkers (those not posting their opinion). Hennig-Thurau et al. (2004) investigated motives of consumer online articulation and proposed that consumers' desire for social interaction, desire for economic incentives, their concern for other consumers, and the potential to enhance their self-worth are the primary factors leading to electronic word of mouth behavior. Toubia and Stephen (2013) empirically studied the motivations of users to contribute content to social media and indicated that image-related utility (the management of the user's image which refers to how the user is portrayed on the platform and the sense of self-worth and social acceptance) is larger than intrinsic utility for most users. Lovett et al. (2013) argued that consumers spread the word on brands as a result of three drivers: functional, social, and emotional. They found that social and functional drivers are the most important for online WOM.

To increase consumers' propensity to engage in online opinion expression, firms have employed many different forms of incentives, including the extrinsic monetary rewards such as promotional payments and nonmonetary rewards such as representations that recognize the expert users. Accordingly, the effect of the employment of those incentives on the consumers' contribution to review websites has also been studied by a number of prior studies.

Sun, Dong, and Mclntyre (2017) revealed an overall decrease in total contributions after introducing monetary rewards for posting reviews. They examined the possible moderating effect of social connectedness (measured as the number of friends) on publicly offered monetary rewards and showed that more-connected members contribute more often when the community relies purely on intrinsic motivation. Khern-am-nuai, Kannan, and Ghasemkhani (2018) compared the quantity and quality of reviews before and after rewards are introduced and found that reviews are significantly more positive, but that quality decreases after rewards are introduced. They also revealed that despite an increase in the number of new reviewers after the monetary rewards were introduced, disproportionately more reviews appear to be written for highly rated products.

This paper empirically identifies systematic biases in online consumer product reviews in rating environments. I examine self-selection bias in the consumers' decision to participate in online conversation by identifying if disproportionately more reviews are written in specific rating environments. When consumers respond to previously posted reviews, this may either increase or decrease the consumer's willingness to write reviews, which is referred to as a selection effect.

My choice of restaurants as the product category for observing the link between previous and subsequent reviews offers a significant advantage. Word of mouth can especially influence intangible products such as service sector like restaurants since customers may not have enough experience before they purchase the service (Zhang et al. 2010; Klein, 1998). The impact of the reviews for these intangible products on potential consumers is significantly huge compared to tangible products.

Using a large dataset consisting of restaurant reviews from Yelp.com, I present the following findings of the impact of previous rating environments on subsequent review generation: (1) more reviews are contributed toward the restaurants with more reviews in the previous period, (2) activists contribute more reviews toward the restaurants with fewer cumulative number of reviews in the previous period, (3) more reviews are contributed toward the restaurants with higher Yelp rating in the previous period, and (4) more reviews are contributed toward the restaurants with a shorter average length of reviews in the previous period. Overall, these results show that disproportionately more reviews are written in the particular rating environments with a consistent pattern of reviewers responding to previously posted reviews.

2. LITERATURE REVIEW

Previous literature has widely studied the causal impact of word of mouth on product adoption and sales (Chevalier and Mayzlin 2006; Liu 2006; Godes and Mayzlin 2009; Luca 2011; Moe and Trusov 2011; Stephan and Galak 2012; Srinivasan et al. 2015). The series of prior research revealed the direct effects of consumer activity in online media on firms’ sales and revenue.

Social media environments have long been of interests to marketing researchers not only to understand consumer behavior but also to explore firms’ marketing strategies since companies effectively leveraged social media as their marketing tool. Dellarocas and Narayan (2006) suggest substantial advances in the ability of organizations to manage word of mouth with the new ability to measure aspects of word of mouth in real-time by mining publically available data from internet communities.

Social dynamics in online review websites have also been extensively investigated. They indirectly affect future product sales in the sense that previously posted reviews have an impact on the subsequent online word of mouth, consequently affecting the future consumer purchases. Social dynamics in social media environments indicate the effect of prior online word of mouth on both formation and expression of subsequent online opinions.

Online opinion behavior in review websites involves two different phases: opinion formation and opinion expression [3]. In the formation phase, consumers understand and evaluate the product right after their purchase. In the expression phase, they make a decision to express their opinions online and participate in the online conversation. This opinion expression phase is in line with a review generation process.

My focus in this paper is on examining systematic biases in review generation process. Factors affecting the consumers' propensity to engage in subsequent ratings have potentials to systematically alter the composition of online reviews. I identify the rating environments that induce the incidence decisions of writing disproportionately more reviews for restaurants in Yelp.com by measuring the relationship between previously expressed opinions and the subsequent review generating process.

[3] See Berinsky (2004) for further explanation of public opinion and political participation in America.

2.1. Social dynamics in the opinion formation phase

Li and Hitt (2008) studied the effect of online word of mouth on the consumers' opinion formation phase. They argued that consumer reviews posted in early periods are systematically positively biased because consumers with higher evaluations tend to purchase and review products first. Nevertheless, consumers do not discount early reviews when they refer to consumer reviews for quality information. This finding provides a rationale for the downward trend over time in online ratings in the sense that late adopters are less satisfied with their purchase decisions based on the evaluation of innovators who may hold very different preferences (product life cycle effect).

Godes and Silva (2012) investigated the impact of online social environment on the consumers' opinion formation phase by examining the evolution of online ratings over time and sequence. They found support for the idea that one's ability to assess the diagnosticity of previous reviews decreases: when previous reviews are very different, more reviews may thus lead to more purchase errors and lower ratings. This finding also explains the downward trend over time in online ratings in the sense that consumers have more difficulty sifting through the posted ratings as the number of ratings available increases (preference matching effects).

Zhao et al. (2013) is another paper that examined the effect of prior online opinions on the opinion formation phase. They modeled consumer learning on both product quality and review credibility by extending the Bayesian learning framework. They found that consumers learn more from online reviews than from their own experience with similar products. Furthermore, they illustrated how the profit impact of product reviews varies with the number of reviews, suggesting the significant effect the number of reviews has on both consumers and firms.

Wu et al. (2015) extended the Zhao et al. (2013) by allowing consumers to learn over multiple attributes (cost and quality) from reading online reviews. They showed that consumers learn their own preferences for multiple product attributes and update not only the expectation but also the variance of their preferences.

2.2. Social dynamics in the opinion expression phase

Moe and Trusov (2011) investigated the effects that previous ratings have on the consumers' opinion expression phase. They modeled the arrival of subsequent ratings within each star level as five separate hazard processes and demonstrated the expected rating behavior both with and without the effect of social dynamics. By separating the effects of social dynamics on ratings from the underlying baseline rating behavior, they measured the effects that previously posted ratings have on future rating behavior. They indicated that increases in average ratings tend to encourage the subsequent posting of negative ratings (one, two, and three-star ratings) and discourage the posting of extremely positive ratings (five-star rating). Furthermore, they showed that disagreement among raters tends to discourage the posting of extreme opinions by subsequent raters.

Moe and Schweidel (2012) revealed a J-shaped relationship between frequency of posts and satisfaction with the product, which is one of the most robust findings in an online word of mouth. The J-shaped relationship suggests that while those with negative opinions are more likely to share an opinion than those with a moderate opinion, those with positive opinions are even more likely to share online (Wierenga 2008). They explored social dynamics affecting the individual-level decisions of whether to post an opinion or what to post by examining the effects of the previously posted content on both posting incidence and evaluation decisions. In terms of incidence decisions, they showed that positive environments increase posting incidence, whereas negative environments discourage posting. Regarding evaluation decisions, they found out that less frequent posters are more positive and exhibit bandwagon behavior, whereas more active posters are more negative and exhibit differentiation behavior. This finding gives another explanation for the downward trend in the sense that the low-involvement groups begin to withdraw from the conversation and refrain from sharing their opinions whereas activists enjoy the dissentious environments and contribute critical opinions. The minority activists who tend to be negative dominate over time, resulting in the downward trend in posted opinions.

Guo and Zhou (2016) showed that both volume and variance of prior ratings exert a negative moderating effect on the relationship between the average rating of prior reviews and the subsequent rating. They found out that such moderating effects are contingent on subsequent reviewer connectedness and expertise. Specifically, when subsequent reviewers become more connected and expert, the effect of the volume of prior ratings will become weaker while the effect of the variance of prior ratings will become insignificant.

Another research stream focused on the consumption value of word of mouth, which is the subsequent phase of the opinion expression. Mudambi and Schuff (2010) indicated that review extremity (measured by the star rating of the review), review depth (measured by the number of words of the review), and product type affect the perceived helpfulness of the review. They showed that moderate reviews are more helpful than extreme reviews for experience goods, but not for search goods. They also revealed that lengthier reviews increase the helpfulness of the review. Pan and Zhang (2011) also showed that review valence and length have positive effects on review helpfulness, but the product type moderates these effects. These studies revealed that the review valence, review extremity, and review length are all important elements that affect the perceived value of consumer-created content.

Whereas prior research has addressed the link between previously contributed opinions and the subsequent opinions based on an individual level or rating level analysis, little work has examined the underpinning of such impact in the aggregate level. The objective of this research is to examine the systematic link between prior and posterior reviews in the aggregate level and reveal the factors that affect the consumers' propensity to engage in ratings in the subsequent period by studying online reviews at the content level (Chevalier and Mayzlin 2006; Luca 2011).

3. DATA

I conduct empirical analysis using the large dataset collected from Yelp.com[4]. I gather reviews and tips[5] for restaurants with at least 500 reviews contributed from January 2005 to November 2018. This leads me to eliminate 5,583,271 reviews, which are 83 percent of the initial 6,685,900 reviews, and 1,007,173 tips, which are 82 percent of the initial 1,223,094 tips, resulting in a data set of 1,102,629 reviews written by 495,537 unique reviewers and 215,921 tips left by 95,888 unique reviewers. Of these reviews, 19 percent were written by elite members, and 60 percent were written by activists. Of the tips, 16 percent were left by elite members, and 51 percent were left by activists. I exclude from the analysis reviews and tips of restaurants without at least a review contributed in the last month of the period of interest, resulting in 1,015,248 reviews and 195,892 tips for 999 restaurants.

The final dataset is at the restaurant month level, consisting of 100,952

---

[4] Yelp provides various tools, including Yelp open dataset for developers(https://www.yelp.com/developers).

[5] For each restaurant, consumers can provide a detailed text describing their restaurant experience along with an overall rating using a discrete five-star scale or write a short text to give information about the restaurant to potential users without a numerical star rating. In this paper, the term 'review' refers to a longer text with a numerical star rating. On the other hand, the term 'tip' refers to a shorter text without a star rating.

observations. Table 1 and Table 2 present the descriptive statistics and variable correlations for the main variables in the final data. The mean rating is 3.96 stars out of 5. On average, a restaurant in my data set receives 10 reviews per month. Of these reviews, 6 reviews come from activists, and 1.9 reviews come from elite members.

<Table 1> Descriptive statistics

Variables	Obs	Mean	SD	Min	Max
1.NumReview	100,952	10.06	11.48	0	247
2.CumNumReview	100,952	424.2	530.34	1	8570
3.AverageRating	100,952	3.96	0.47	1	5
4.YelpRating	100,952	3.84	0.51	1	5
5.VarRating	100,952	1.21	0.46	0	8
6.AverageLength	100,952	133.1	39.11	2	926

Note: All statistics are per month per restaurant

<Table 2> Variable correlations

Variables	1	2	3	4	5	6
1.NumReview	1					
2.CumNumReview	0.61	1				
3.AverageRating	0.08	0.00	1			
4.YelpRating	0.08	0.00	0.95	1		
5.VarRating	0.14	0.20	-0.57	-0.55	1	
6.AverageLength	-0.09	-0.14	-0.12	-0.12	-0.03	1

Note: All statistics are per month per restaurant

I employ two different measures to characterize reviewers following Guo and Zhou (2016). First, I measure reviewer involvement using the total number of reviews the reviewer had posted during the period of interest (Liu and Park 2015; Racherla and Friske 2012; Guo and Zhou 2016). Reviewers in Yelp.com write 22 reviews on average during the entire period. I consider a reviewer an activist if the cumulative number of reviews contributed by the reviewer exceeds the average number of reviews that Yelp reviewers have contributed.

Second, I measure reviewer expertise based on whether the reviewer is a Yelp Elite member (Chen and Lurie 2013). Yelp recognizes people who are active in the Yelp community and role models on and off the site. They designate an Elite badge on the members' profile based on well-written reviews, high-quality tips, a detailed personal profile, an active voting and complimenting record, and a history of playing well with others. The Yelp Elite Squad is a yearly program, so badges will only extend until the end of the calendar year.[6]

[6] See https://www.yelp-support.com/article/What-is-Yelps-Elite-Squad?l=en_US for a more detailed explanation on What Yelp's Elite Squad is.

4. HYPOTHESIS DEVELOPMENT

A series of previous research has revealed social dynamics in the opinion expression phase. Hu et al. (2006) link biases in review platforms with the consumers' motivation of leaving a review online. They note the reviewers' bragging and moaning behavior that they only choose to write reviews when they are very satisfied with the products they purchased (brag), or very disgruntled (moan). This behavior is in a similar vein with a J-shaped relationship between frequency of posts and satisfaction with the product (Moe and Schweidel 2012), which is one of the most robust findings in an online word of mouth.

Overall, these studies focus on the link between the product experience and the incidence decision. Whereas prior research has addressed the relationship between post purchase evaluation and the incidence decision, little work has examined the underpinning of the link between the posted product ratings and the subsequent incidence decision of expressing opinions.

In this paper, I develop hypotheses about the relationship between the rating environments and the incidence decision of engaging in writing reviews. The incidence with which individuals choose to express their own opinions is closely related to the motivation and the incentive of leaving an opinion on the review websites. If more reviews appear to be written for restaurants with a certain rating environment, particular rating environments might have incentivized more consumers to express opinions. The tendency toward posting more reviews in the specific rating environments implies a bias in selecting restaurants to leave a review.

Volume (Liu 2006; Duan et al. 2008; Wang 2015), valence (Liu 2006; Duan et al. 2008; Chevalier and Mayzlin 2006; Wang 2015), and variance (Godes and Mayzlin 2004; Chintagunta et al. 2010) are the three common measurements of product ratings that have been primarily used in the prior literature. I follow previous research and examine these measurements to characterize the rating environments of the online review website. Additionally, I include the average length (Ma et al. 2013) in the measurements.

4.1. The Effect of Review Volume on Incidence Decisions

Review volume measured by the total number of reviews captures the frequency of the consumers' opinion expression on the entity. The low frequency of the consumers' contribution in review websites has long been of interests to considerable research and prompted the researchers to investigate what motivates consumers to express opinions on the review websites.

Recall that a series of previous research has revealed the intentions behind why individuals express opinions publically: social, emotional, and functional (Lovett et al. 2013). The social driver relates to social signallings such as expressing uniqueness (the management of the image), self-worth, self-enhancement, social acceptance, and a desire for social interaction. The emotional driver refers to emotion sharing. The functional driver is related to the tendency to exchange useful and practical information.

The social drivers of word of mouth include not only the behavior of expressing uniqueness but also the behavior of tendency towards conformity. The bandwagon effect, which refers to the probability of individual adoption increasing with respect to the proportion who have already done, suggests that the opinions of others can influence an individual's decision of whether or not to express an opinion.

Consumers occur the cumulative number of reviews when searching for a restaurant in Yelp.com.[7] The aggregate number of reviews representing

---

[7] See Figure 1 to identify how users in Yelp.com are informed of the cumulative number of reviews written for the specific restaurant.

how many others have already written reviews for the restaurant might have affected an individual's propensity to engage in subsequent ratings in the sense that the bandwagon effect arises when the consumers are noticed that others have already engaged in the same behavior.

Zhao et al. (2013) revealed how the profit impact of product reviews varies with the number of reviews, implying the significant effect the number of reviews has on consumers who seek information through reading reviews before making their purchases. The theory behind measuring volume, or the number of online messages posted on a topic, is that the more consumers discuss a product, the higher the chance that other consumers will become aware of it (Dellarocas et al. 2006).

Furthermore, a study published in psychological science found that customers are more likely to favor a product based on the quantity of reviews, rather than what they say. The findings indicate that most people tend to prefer a product that has more reviews, even if it has a lower rating than another product, suggesting that the number of reviews is an indication of a business's credibility and reputation.[8]

---

[8] https://business.trustpilot.com/reviews/more-is-better-why-review-quantity-matters (accessed June 14, 2019). See Powell et al. (2017) for more specific results of the study.

The previous studies reveal that review quantity affects consumer behavior in many different ways and prompt the idea that the numerical number of reviews for a specific product or service might have a direct impact on the incidence decisions through prompting the herd behavior as well as an indirect impact on the incidence decisions through affecting the consumers' purchase decisions. In other words, consumers are more likely to write reviews for the restaurants with a large number of reviews not only because they tend to follow the behavior of others but because they have a higher chance of going to those restaurants since they believe that the restaurants are highly regarded by others.

Based on the above discussion, the overall relationship between the number of reviews contributed in the previous period and the consumers' incidence decisions of expressing opinions in the subsequent period is not obvious. Therefore, I motivate the following hypotheses as empirical questions:

Hypothesis 1A. More reviews are contributed toward the restaurants with more reviews in the previous period.

Hypothesis 1B. More reviews are contributed toward the restaurants

with the more cumulative number of reviews in the previous period.

The tendency toward engaging in online conversation for the restaurants that already have a number of reviews can be explained by the emotional driver in the sense that people tend to recount and share emotional experiences with others.

Guo and Zhou (2016) revealed that the volume of prior ratings exerts a negative moderating effect on the relationship between the average rating of prior reviews and the subsequent rating. They found out that such moderating effects are contingent on subsequent reviewer expertise. When subsequent reviewers become more expert, the effect of the volume of prior ratings will become weaker. This finding implies that the expert reviewers might respond less to the volume of prior ratings than overall reviewers.

Moe and Schweidel (2012) found out that less frequent posters are more positive and exhibit bandwagon behavior in the sense that they are easily influenced by others, whereas more active posters are more negative and exhibit differentiation behavior. The differentiation behavior of active posters suggested by the previous research can be identified by the following empirical questions:

Hypothesis 1C. Activists contribute more reviews toward the restaurants with fewer reviews in the previous period.

Hypothesis 1D. Activists contribute more reviews toward the restaurants with the fewer cumulative number of reviews in the previous period.

4.2. The Effect of Review Valence on Incidence Decisions

The theory behind a valence representing the fraction of positive and negative opinions in the online messages is that word of mouth carries important information about a product’s quality (Dellarocas et al. 2006). Dellarocas et al. (2005) found that the valence of online ratings posted during a movie’s opening weekend was the most significant predictor of that movie’s revenue in subsequent weeks.

Moe and Schweidel (2012) revealed a J-shaped relationship between the frequency of posts and satisfaction with the product, showing the link between product evaluation and the incidence decisions. The positivity bias, which indicates that the online word-of-mouth tends to be predominantly positive, has also been demonstrated across previous studies (Anderson 1998, Resnik and Zeckhauser 2002, Chevalier and Mayzlin 2006, Godes and Mayzlin 2004). These studies examined that consumers tend to express opinions more positively compared to their actual level of satisfaction.

However, little work has examined the link between the posted review valence and the subsequent incidence decision of expressing opinions. I examine the relationship between previous ratings and the incidence decisions by identifying if more reviews are disproportionately written for highly rated restaurants. I have the following two empirical hypotheses:

Hypothesis 2A. More reviews are contributed toward the restaurants with a higher actual rating in the previous period.

Hypothesis 2B. More reviews are contributed toward the restaurants with a higher Yelp rating in the previous period.

The tendency toward writing more reviews for the restaurants that have higher rating can be explained by the social driver in the sense that the behavior of writing reviews for the restaurants that have already been highly recognized by other consumers can give the sense of self-worth and social acceptance.

4.3. The Effect of Review Variance on Incidence Decisions

Godes and Silva (2012) found that when previous reviews are very different, more reviews may thus lead to more purchase errors and lower ratings. Moe and Trusov (2011) showed that disagreement among raters tends to discourage the posting of extreme opinions by subsequent raters. Sun (2012) revealed that a high variance of ratings is associated with a niche product and a higher variance would correspond to a higher subsequent demand if and only if the average rating is low. These studies have revealed the impact of disagreement among prior reviewers on consumers' purchase and the incidence decisions or posting reviews.

I further investigate the relationship between the variance of prior ratings and the consumers' subsequent decisions of whether or not expressing opinions through the following empirical question:

Hypothesis 3. More reviews are contributed toward the restaurants with

a lower opinion variance in the previous period.

4.4. The Effect of Review Length on Incidence Decisions

Mudambi and Schuff (2010) measured review depth by the number of words in the review and indicated that review depth affects the perceived helpfulness of the review. Pan and Zhang (2011) also showed that the review length has positive effects on review helpfulness. These studies revealed that the review length is an important element that affects the perceived value of word of mouth. However, little work has examined how the review length is related to the subsequent incidence decisions of writing reviews.

In this paper, I assume that having many words in previously contributed reviews infers that the overall prior opinions contain more information. The functional driver, which is related to the tendency to exchange useful and practical information, can prompt consumers to write more reviews for the restaurants with a shorter average length of prior reviews which do not give enough information for potential consumers at that point. This motivates the following empirical question:

Hypothesis 4. More reviews are contributed toward the restaurants with a shorter average length of reviews in the previous period.

5. EMPIRICAL ANALYSIS AND RESULT

Regression analysis is well suited for identifying the above empirical questions since it allows me to examine the relationships between main variables that are of interest to this research. I model the total number of reviews in the subsequent period as a linear combination of variables that indicate the rating environments in the previous period to investigate how previously posted reviews are related to the total number of reviews that are subsequently posted. It is assumed that there is a significant relationship between the previous rating environments, which are identified by the measures of volume, valence, variance, and length, and the subsequent number of reviews to be posted.

The dependent variable, $NumReviews_{it}$, is the total number of reviews, which indicates the volume of opinion expression for a specific restaurant i written during the time period t.

Independent variables are as follows. $NumReviews_{it-1}$(total number of reviews contributed in the previous period t-1 for a restaurant i) and

$CumNumReviews_{it-1}$(Cumulative number of reviews for a restaurant i as of the previous time period t-1) are the measurements of review volume.

$AverageRating_{it-1}$(Cumulative numerical average ratings score for a restaurant i as of the previous time period t-1) and $YelpRating_{it-1}$ (Cumulative rounded average rating score displayed on Yelp.com for a restaurant i as of the previous period t-1) are the measurements of review valence. Since the star rating displayed on the Yelp.com is not the exact average star rating as shown in Figure 1, I calculate not only the cumulative actual overall average rating of restaurants ($AverageRating_{it-1}$) but also cumulative rounded overall average rating (rounded to the nearest half-star) displayed by Yelp.com ($YelpRating_{it-1}$).

$VarRating_{it-1}$(Cumulative variance in ratings for a restaurant i as of the previous time period t-1) is a measurement of review variance. $AverageLength_{it-1}$(Cumulative numerical average number of words for reviews of a restaurant i as of the previous time period t-1) is a measurement of review length.

I control unobserved heterogeneity among different restaurants (α_i) by using fixed effects dummy variables. The restaurant fixed effect is related to factors such as food quality (taste, the freshness of meals, and amount of food), hygiene (clean dining area and clean staff), responsiveness (prompt service) and menu (display, variety, and knowledge of items) (Almohaimmeed 2017).

I account for time heterogeneity (δ_t) by using dummy variables to control for any monthly time effects that are constant across entities but vary over time. As a result, I rule out environmental factors such as time trends that can influence the subsequent review volume.

I control for all of the other factors on the rating page that can affect the number of reviews in the subsequent period. I represent the vector of control variables as X_{it-1}. Control variables include the measure of the consumption value of previously posted reviews by using the cumulative numerical average number of votes (useful, funny, and cool) that the restaurants had received ($NumVotes_{it-1}$).[9] I also control for the rating effects that are related to the reviewer characteristics by using the volume, valence, variance, and length of reviews contributed by Elite members and activists ($EliteNumReviews_{it-1}$, $ActivistNumReviews_{it-1}$, $EliteCumNumReview_{it-1}$, $ActivistCumNumReviews_{it-1}$, $EliteRating_{t-1}$,

$ActivistRating_{t-1}$, $EliteVarRating_{t-1}$, $ActivicstVarRating_{t-1}$,

[9] See Figure 2 to identify how users in Yelp.com vote for the review for the restaurant while they are scrolling down the page to read reviews.

$EliteAverageLength_{it-1}$, $ActivistAverageLength_{it-1}$). I account for any other rating effects such as the cumulative variance in the average length of reviews ($VarLength_{it-1}$). I also account for the cumulative numerical average number of tips written for each restaurant in the previous time period ($NumTip_{it-1}$).

For restaurant i at time t, my main model specification to measure the incidence decision of the overall users in Yelp.com is as follows:

$$
\begin{aligned}
NumReview_{it} = \alpha_i + \delta_t + \beta X_{it-1} + \gamma NumReviews_{it-1} \\
+ \theta CumNumReviews_{it-1} + \eta AverageRating_{it-1} \\
+ \zeta YelpRating_{it-1} + \mu VarRating_{it-1} + \rho AverageLength_{it-1} \\
+ \psi AverageLength_{it-1}^{2} \\
+ \varepsilon_{it} \qquad (1)
\end{aligned}
$$

The models to be tested for identifying the incidence decision of elite members and activists are as follows:

$$\begin{aligned}
\mathrm{EliteNumReview}_{\mathrm{it}} &= \alpha_{\mathrm{i}} + \delta_{\mathrm{t}} + \beta \mathrm{X}_{\mathrm{it}} + \gamma \mathrm{NumReviews}_{\mathrm{it-1}} + \theta \mathrm{CumNumReviews}_{\mathrm{it-1}} \\
&+ \eta \mathrm{AverageRating}_{\mathrm{it-1}} + \zeta \mathrm{YelpRating}_{\mathrm{it-1}} + \mu VarRating_{it-1} \\
&+ \rho \mathrm{AverageLength}_{\mathrm{it-1}} + \psi \mathrm{AverageLength}^{2}_{\mathrm{it-1}} \\
&+ \varepsilon_{\mathrm{it}} \qquad (2)
\end{aligned}$$

$$\begin{aligned}
\mathrm{ActivistNumReview}_{\mathrm{it}} &= \alpha_{\mathrm{i}} + \delta_{\mathrm{t}} + \beta \mathrm{X}_{\mathrm{it}} + \gamma \mathrm{NumReviews}_{\mathrm{it-1}} + \theta \mathrm{CumNumReviews}_{\mathrm{it-1}} \\
&+ \eta \mathrm{AverageRating}_{\mathrm{it-1}} + \zeta \mathrm{YelpRating}_{\mathrm{it-1}} + \mu VarRating_{it-1} \\
&+ \rho \mathrm{AverageLength}_{\mathrm{it-1}} + \psi \mathrm{AverageLength}^{2}_{\mathrm{it-1}} \\
&+ \varepsilon_{\mathrm{it}} \qquad (3)
\end{aligned}$$

Table 3 reports result from the above fixed effect models.

<Table 3> Results from the Regression Analysis

	Number of Reviews from overall users	Number of Reviews from Elite members	Number of Reviews from activists
Number of Reviews	0.546*** (0.005)	0.030*** (0.002)	0.166 *** (0.004)
Cumulative	0.000 (0.012)	0.001*** (0.000)	-0.003 ***(0.000)

Number of Reviews			
Average Rating	-0.041 (0.317)	-0.184. (0.109)	-0.127 (0.214)
Yelp Rating	0.506 *** (0.118)	0.184 *** (0.041)	0.365 ***(0.080)
Variance of Ratings	-0.184 (0.153)	-0.086 (0.052)	-0.042 (0.103)
Average Length	-0.041 *** (0.007)	-0.002 (0.002)	-0.028 *** (0.004)
Average Length ×Average Length	0.0001*** (0.000)	0.000 (0.000)	0.000*** (0.000)
Restaurant fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Observations	97,369	97,369	97,369
Adjusted R-squared	0.9032	0.7664	0.8795

Note: *** indicates that the p-value is less than 0.001.

The model fits the data very well: Adjusted R-squared is 0.9032, and the p-value is less than 2.2e-16, indicating that the regression equation explains 90% of the variation in the dependent variable.

The estimate for γ, the response to the number of prior reviews, is 0.546, statistically significant at the 0.001 level. This suggests that the volume of reviews of a restaurant in the previous period can significantly increase the incidence decision of consumers to participate in the rating, indicating that the Hypothesis 1A is supported empirically. However, the response to the cumulative number of prior reviews is not statistically significant for overall users, suggesting that the cumulative number of previously posted reviews is statistically not related to the number of reviews posted in the subsequent period.

In contrast to the results from the regression analysis for the overall reviewers, the activists' response to the cumulative number of prior reviews is negative and statistically significant, suggesting that Hypothesis 1D is supported. In contrast to the activists' response, the elite members' response to the cumulative number of reviews in the previous period is positive and statistically significant.

What is especially interesting in my finding is that the incidence decisions of consumers significantly differ depending on the reviewers' activity level and status in the review website. The activists' behavior of writing more reviews toward the restaurants with the fewer cumulative number of reviews can be explained as a differentiation behavior. The differentiation behavior is usually prompted by the social drivers of word of mouth as activists are motivated to express uniqueness a lot more than overall users in Yelp.com.

Most importantly, I find that Yelp rating (valence) has a positive and statistically significant impact on the number of reviews posted subsequently, indicating that the data supports Hypothesis 2B. This finding is explained in the sense that users who are strictly rational utility maximizers can minimize their effort by choosing highly rated products and borrowing product specifications and characteristics already mentioned in the existing reviews (Khernamnuai et al. 2018). However, the parameter estimate for the actual Average Rating is not statistically significant, suggesting that consumers largely rely on the star rating displayed by Yelp rather than calculating the average rating of the restaurant by themselves.

The coefficient of Variance of Ratings is not statistically significant, indicating that the opinion variance is statistically not correlated to the incidence decisions in the subsequent period. This result is inconsistent with empirical findings in the literature that provide a significant relationship between an individual's decision of providing a rating and whether or not an agreement or a dissention exists in the rating environments (Moe and Schweidel 2012).

The parameter estimate for the quadratic terms of Average Length is positive and statistically significant, suggesting a nonlinear relationship between Average Length and the Number of Reviews. This result reveals that the average review length plays a different role in prompting consumers to post reviews depending on the level of the review length.

A possible reason for the increased incidence decisions when the average review length is short is that consumers are motivated to contribute more when they sense that the previously posted reviews lack information. The functional driver might incentivize them to engage in online word of mouth for providing useful information that has not been available to the users. At the same time, consumers are motivated to participate in a conversation more when a lively discussion is already going on. This tendency might have been facilitated by social drivers: a desire for social interactions and sharing opinions. The social driver explains the increased engagements in ratings when the average review length for the restaurants is short.

To summarize, the total number of reviews, Yelp rating, and the average length of reviews written in the previous period are all significant predictors of the number of reviews posted in the subsequent period, which indicates consumers' propensity to engage in subsequent ratings.

Table 4 summarizes my hypotheses and results.

<Table 4> Summary of Results of the Hypothesis Tests

Measure		Hypotheses	Results
Volume	1A	More reviews are contributed toward the restaurants with more reviews in the previous period.	Supported
	1B	More reviews are contributed toward the	Not

		restaurants with more cumulative reviews in the previous period.	supported
	1C	Activists contribute more reviews toward the restaurants with fewer reviews in the previous period.	Not supported
	1D	Activists contribute more reviews toward the restaurants with fewer cumulative reviews in the previous period.	Supported
Valence	2A	More reviews are contributed toward the restaurants with a higher actual rating in the previous period.	Not Supported
	2B	More reviews are contributed toward the restaurants with a higher Yelp rating in the previous period.	Supported
Variance	3	More reviews are contributed toward the restaurants with a lower opinion variance in the previous period.	Not supported
Length	4	More reviews are contributed toward the restaurants with a shorter average length of reviews in the previous period.	Supported

6. CONCLUSION

This study has developed a regression model to capture the relationship between the previous rating environments and the number of reviews that will be posted subsequently. Through the regression analysis, I find that there exist systematic biases in the volume of online consumer reviews after controlling for restaurant and time fixed effects.

I focus on the restaurant reviews and use the data collected from Yelp.com to show that previously posted reviews exhibit a number of empirical relationships with consumers' propensity to engage in online opinion expression in the subsequent period.

Several novel findings emerge from my analysis. First, I observe that individuals are more prone to post reviews for the restaurants with more reviews contributed in the previous period. They exhibit bandwagon effects by adjusting their incidence decisions when more people have already contributed to the ratings previously. Whereas activists provide more reviews toward the restaurants with a fewer cumulative number of reviews, overall users do not respond to the cumulative number of reviews. The result indicates noticeably different posting behavior between active posters and overall users.

I also observe an increased likelihood of posting incidence for highly rated restaurants, consistent with previous research showing that consumers are more likely to post an opinion when the ratings already posted are more positive (Moe and Schweidel 2012). Particularly, I find that consumers' posting behavior does not respond to the actual average rating of the restaurants, suggesting the relative significance of Yelp rating in the consumers' incidence decisions. This finding is especially important given that more and more companies are putting social media at the core of their marketing strategy to attract the attention of consumers and give consumers a sense of credibility for the products and services.

Furthermore, I see that individuals are more likely to write reviews when the average length of previous reviews is either very short or very long. This empirical finding illustrates the potentially nonlinear relationship between the subsequent incidence decision and the information availability in the previous rating environments. The finding can be used to draw insights on incentive designs for increasing the level of participation in the review platforms.

I note that the variance of ratings does not have a significant impact on the incidence decisions. This result is not consistent with the previous literature that revealed the considerable effect of opinion variance on the subsequent opinion expression. Therefore, further work is required to establish whether the opinion variance significantly affects the consumers' propensity to participate in the online word of mouth in different contexts.

Overall, online reviews are disproportionately written for certain rating environments with a consistent pattern of consumers responding to previously posted reviews. My findings represent an important extension to previous research on the relationship between the specific rating environment and consumers' propensity to engage in subsequent ratings. This finding is particularly important as it highlights that the review volume, review valence, and review length can all serve as a useful proxy of an aggregate number of reviews that will be posted in the subsequent period.

My analysis has several limitations that suggest directions for future research. First, the correlation between the rating environments in the previous period and the number of reviews contributed subsequently does not necessarily imply causality. Thus, a future research direction would be to understand the causality by conducting field experiments or laboratory experiments.

Second, my analysis focuses on the restaurant category. However, considerable heterogeneity across different categories is expected, shedding light on the conditions under which previous reviews affect an individual's propensity to contribute to online conversation differently in the categories other than the restaurant category. Therefore, another avenue for future research is to explore whether the effect of rating environments is similar or different for the reviews of other product categories.

More broadly, my results relate to an emerging discussion about the mechanism through which social dynamics in the opinion expression phase have significant associations with subsequent reviewers' incidence decisions. The results not only draw scrutiny from firms for understanding and leveraging social media, and managing their marketing tools but also alert consumers to recognize their bias in selecting products to write reviews.

REFERENCES

Almohaimmeed (2017) Restaurant Quality and Customer Satisfaction. International Review of Management and Marketing 7(3), 42-49.

B. Wierenga (2008) Handbook of Marketing Decision Models: Chapter 16 Social Media Analytics, 483-504.

Berinsky, A.J. (2004) Silent Voices: Public Opinion and Political Participation in America. Princeton University Press, Princeton.

Chen Z, Lurie N.H. (2013) Temporal contiguity and negativity bias in the impact of online word-of-mouth. J. Mark. Res. 50(4).

Chevalier, Judith and Dina Mayzlin (2006) The Effect of Word-of-Mouth on Sales: Online Book Reviews. Journal of Marketing Research Vol. 43 No 3.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. Marketing Sci. 29(5).

Dellarocas C, Awad N, Zhang M (2005) Using online ratings as a proxy of word-of-mouth in motion picture revenue forecasting. Smith School of Business, Univ. Maryland.

Dellarocas C, Narayan R (2006) A statistical measure of a populations propensity to engage in post-purchase online word-of-mouth. Statist. Sci. 21(2).

Duan W, Gu B, Whinston, A.B. (2008) The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry. J. Retail. 84(2).

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. Marketing Science. 23(4).

Godes D, Mayzlin D (2009) Firm created word of mouth communication: Evidence from a field test. Marketing Science. 28(4).

Godes D, Silva J (2012) Sequential and temporal dynamics of online opinion. Marketing Science. 31(3).

Guo Bin, Zhou Shasha (2016) Understanding the impact of prior reviews on subsequent reviews: The role of rating volume, variance and reviewer characteristics. Electronic Commerce Research and Applications 20.

Hennig-Thurau T, Gwinner KP, Walsh G, Gremler DD (2004) Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? J. Interactive Marketing 18(1).

Hu Y, Li X (2011) Context-dependent product evaluations: an empirical analysis of internet book reviews. J. Interact. Mark. 25(3).

Hu N, Pavlou PA, Zhang J (2006) Can online reviews reveal a product's true quality: Empirical findings and analytical modeling of online word-of-mouth communication. Proc. 7th ACM Conf. Electronic Commerce (ACM, New York).

Hu N, Zhang J, Pavlou PA (2009) Overcoming the j-shaped distribution of product reviews. Comm. ACM 52(10).

Khern-am-nuai W, Kannan K, and Ghasemkhani H (2018) Extrinsic versus intrinsic rewards for contributing reviews in an online platform. Information Systems Research 29(4).

Li X, Hitt LM (2008) Self-selection and information role of online product reviews. Inform. Systems Res. 19(4).

Liu Y (2006) Word of mouth for movies: Its dynamics and impact on box office revenue. J. Marketing 70(3).

Liu Z, Park S (2015) What makes a useful online review? Implication for travel product websites. Tourism Manage. 47.

Lovett MJ, R. Peres, R. Shachar (2013) On brands and word of mouth. Journal of Marketing Research 50 (4).

Luca M (2011) Reviews, reputation, and revenue: The case of Yelp.com. Working paper, Harvard Business School.

Ma X, Khansa L, Deng Y, Kim SS (2013) Impact of prior reviews on the subsequent review process in reputations systems. J. Manage. Inf. Syst. 30 (3).

Moe WW, Schweidel DA (2012) Online product opinions: Incidence, evaluation, and evolution. Marketing Science. 31(3).

Moe WW, Trusov M (2011) The value of social dynamics in online product ratings forums. J. Marketing Res. 48(3).

Mudambi SM and Schuff D (2010) What makes a helpful online review? A study of customer reviews on Amazon.com. MIS Quart. 34(1).

Pan Y, Zhang JQ (2011) Born unequal: A study of the helpfulness of user-generated product reviews. J. Retailing. 87 (4).

Powell D., Yu J., DeWolf M., Holyoak K.J. (2017) The love of large numbers: a popularity bias in consumer choice. Psychological science, 28(10).

Racherla P, Friske W (2012) Perceived usefulness of online consumer reviews: an exploratory investigation across three services categories. Electron. Commer. Res. Appl. 11(6).

Schlosser AE (2005) Posting versus lurking: communicating in a multiple audience context. J. Consum. Res. 32(2).

Stephan AT, J Galak (2012) The effects of traditional and social earned media on sales: A study of a microlending marketplace. Journal of Marketing Research. 49(5).

Srinivasan S, OJ Rutz, K. Pauwels (2015) Paths to and off purchase: Quantifying the impact of traditional marketing and online consumer activity. Journal of the Academy of Marketing Science 1-14.

Sun, M (2012) How does the variance of product ratings matter? Management Science 58(4).

Sun, Dong, Mclntyre (2017) Motivation of User-Generated Content. Marketing Science. Vol 36. No. 3.

Toubia O, AT Stephen (2013) Intrinsic vs image-related utility in social media: Why do people contribute content to twitter? Marketing Science 32(3).

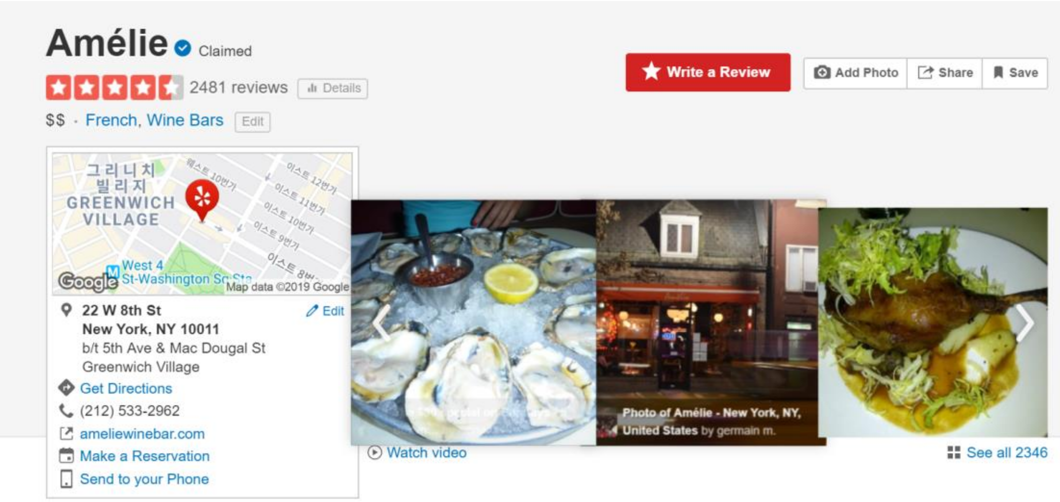
Wang A, Zhang M, Hann IH (2015) Socially nudged: a quasi-experimental study of friends' social influence in online product ratings. Inf. Syst. Res.

Wu C, Che H, Chan T, Lu X (2015) The economic value of online reviews. Marketing Science 34(5).

Zhao Y, Yang S, Narayan V, Zhao Y (2012) Modeling consumer learning from online product reviews. Marketing Science 32(1).

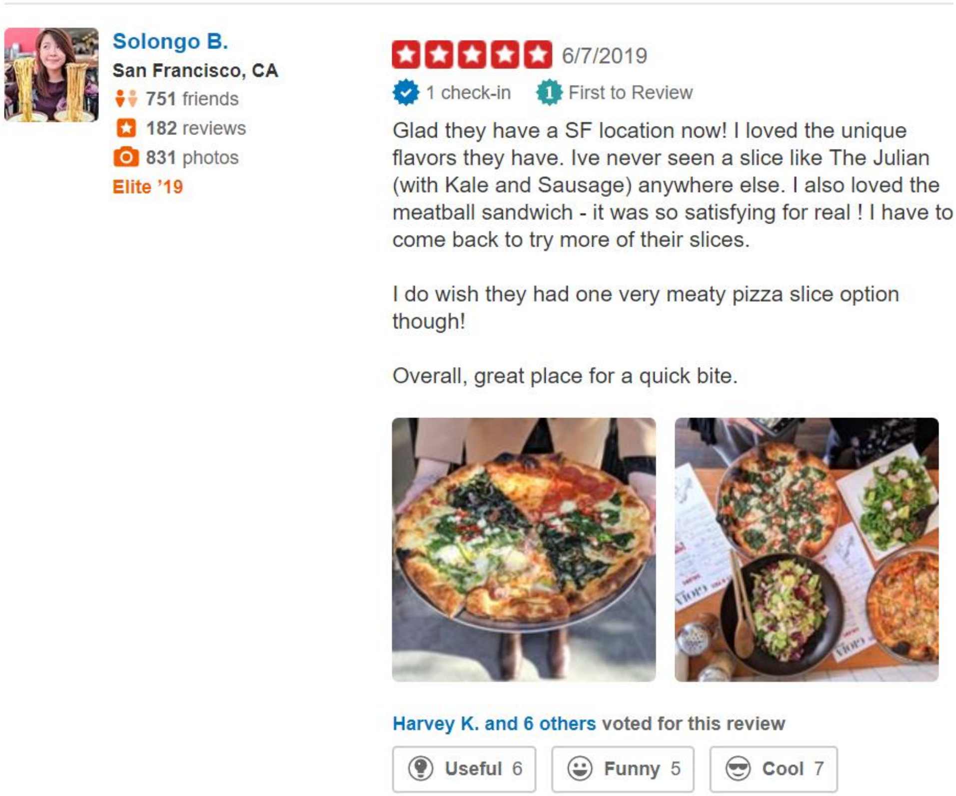
APPENDIX

<Figure 1> sample snapshot of a restaurant in Yelp.com



Note: Yelp Rating and the cumulative number of reviews are displayed on the webpage that consumers first occur when searching a restaurant on Yelp.com.

<Figure 2> sample snapshot of a review in Yelp.com



Note: Users on Yelp.com vote for the review by clicking the useful, funny, and cool buttons displayed right below the written text. Users are informed of various information such as the number of votes the review has received, the number of reviews the reviewer has written, and whether the review was written by an elite member or not.

국문초록

리뷰 작성 환경과 소비자들의 리뷰 작성 경향

김 은 선

경영학과 마케팅전공

서울대학교 대학원

리뷰 작성 환경이 소비자들의 리뷰 작성 경향에 영향을 미칠까? 이전 연구에서 주로 구매 후 평가와 리뷰 작성 경향에 대해 주로 다룬 반면, 작성된 리뷰들과 이후 리뷰 작성 경향에 대한 연구는 많이 이루어지지 않았다. 이 연구에서는 Yelp.com에서 수집한 대용량 리뷰 데이터를 활용하여 의견 표현 단계에서의 Social Dynamics에 대하여 연구한다. 즉, 이전 리뷰와 이후 리뷰들의 체계적인 연결 관계를 분석함으로써 리뷰 작성 경향에 영향을 미치는 사회적인 요소들을 찾아내고자 한다. 이와 같은 요소들은 리뷰 작성 경향에 영향을 미치는 것에서 더 나아가 리뷰 웹사이트 내 리뷰들의 구성 자체를 체계적으로 변경시킬 수 있다. 이 논문은 레스토랑 리뷰 데이터를 분석하여 리뷰 웹사이트 상에 작성된 총 리뷰 수의 체계적인 편향을 경험적으로 확인함으로써 소비자들이 온라인 리뷰 웹사이트 상에 기여할지 여부, 즉 의견 표출 여부를 결정하는 과정에서 발생하는 Self-selection에 대해 검토한다. 연구를 통해 리뷰 작성 환경과 리뷰 작성 경향에 대한 다음의 결과를 보였다. (1) 전체 소비자들은 이전 시점에 작성된 리뷰가 많은 레스토랑에 더 많은 리뷰를 작성한다. (2) 활동적인 소비자들은 누적 리뷰 수가 적은 레스토랑에 더 많은 리뷰를 작성한다. (3) 전체 소비자들은 누적 Yelp 평점이 높은 레스토랑에 더 많은 리뷰를 작성한다. (4) 전체 소비자들은 리뷰의 누적 평균 길이가 짧은 레스토랑에 더 많은 리뷰를 작성한다. 이러한 결과는 특정한 리뷰 작성 환경 상에서 온라인 리뷰가 더 많이 작성되며 소비자들이 기존에 작성된 리뷰에 일관적으로 반응하는 것을 나타낸다.

주요어: 온라인 입소문, 사회적 영향, 리뷰 작성 환경, 리뷰 작성 경향

학 번: 2017-24741