Print ISSN: 1738-3110 / Online ISSN 2093-7717
http://dx.doi.org/10.15722/jds.15.10.201710.41

The Effect of Review Behavior on the Reviewer's Valence in Online Retailing

Yun-Kyung Oh*

Received: August 11, 2017. Revised: August 24, 2017. Accepted: October 15, 2017.

Abstract

Purpose – Online product review has become a crucial part of the online retailer's market performance for a wide range of products. This research aims to investigate how an individual reviewer's review frequency and timing affect her/his average attitude toward products.

Research design, data, and methodology – To conduct reviewer-level analysis, this study uses 42,172 posted online review messages generated by 6,941 identified reviewers for 59 movies released in the South Korea from July 2015 to December 2015. This study adopts Tobit model specification to take into account the censored nature and the selection bias arising from the nature of J-shaped distribution of movie rating.

Results – Our estimation results support that the negative impact of review frequency and timing on valence. Furthermore, review timing has an inverted-U relationship with the user's average valence and enhance the negative effect of review frequency.

Conclusions – This study contributes to the growing literature on the understanding how eWOM is generated at the individual consumer level. On the basis of the main empirical findings, this study provides insights into building a recommendation system in online retail store based on the consumer's review history data - frequency, timing, and valence.

Keywords: Online Product Review, Review Behavior, Review Frequency, Review Timing, Online Retailing.

JEL Classifications: M31, L86.

1. Introduction

With the development of the digital economy, the online review has profoundly influenced consumers' information process, their product choices, and thereby affected a retailer's market performance. The user's experience with a product is disseminating through the Internet with low-cost and mostly available to the public. Nowadays, consumers are increasingly posting their reviews of using a product at an online shopping mall, review site or mobile application. For example, they review products on the online retail site such as Amazon.com, movies on IMDb.com, hotels on Hotels.com, and taxi services on Uber mobile application. Online word-of-mouth increases awareness of a product and provides a reliable information source for the product. As a result, researchers and practitioners have long noticed the importance of online WOM on a product's market performance. Consumers' perception of the product quality can be affected by the publicly available consumer ratings and reviews as well as the product information provided by the seller(Sthapit, Jo, & Hwang, 2016). As a result, strongly positive ratings can positively influence the growth of product sales(Clemons, Gao, & Hitt, 2006).

The online review can help potential buyers to make proper choice decisions with lower information search costs. Chevalier and Mayzlin(2006) examine the effect of consumer reviews on relative sales for online retailers Amazon.com and Barnsandnoble.com. They suggest that favorable reviews at one retailer lead to an increase in the sales of a product at the store relative to another retailer. Yet, most of the prior studies have focused on the aggregated WOM effect on the financial outcome. In contrast, few studies have examined at how individual reviewer level characteristics affect the reviewer's overall valence. Although this study intends to contribute to multiple industries including retail business, we focus on analyzing reviewer behavior in the movie industry. Prior studies on eWOM have concentrated on the role of WOM for experience goods because the quality is hardly known before use. For this reason, the film industry is a popular category for WOM studies. Both researchers and practitioners have believed that WOM affects an individual's movie selection and play a significant role in box office performance(Bayus, 1985; Neelamegham & Chintagunta, 1999; Neelamegham & Jain, 1999). While most of the prior research has conducted at movie-level analysis,

* Assistant Professor, Department of Business Administration, Dongduk Women's University, Seoul, Korea.
Tel: +82-2-940-4471, E-mail: ykoh1@dongduk.ac.kr

the study on understanding reviewer behavior remains an under-investigated area.

Recent studies have examined the role of online customer product reviews, specifically looking at the characteristics of the viewers(Formanm, Ghose, & Wiesenfeld, 2008; Smith, Menon, & Sivakumar, 2005). Other research shows that self-selection bias in rating arises because consumers with different preference join at their preferred period(Hu, Zhang, & Pavlou, 2008; Li & Hitt, 2008). Moon, Bergey, and Iacobucci (2010) conduct an individual viewer-level analysis to examine how individual viewers' movie consumption affects their satisfaction. In this vein, it might be valuable to examine who tend to create more favorable or less favorable reviews. Considering the positive influence of favorable reviews on future sales(Clemons et al., 2006), identifying the characteristics of reviewers that affect their overall valence is an important question.

This study examines the relationship of three primary variables of a focal reviewer using individual reviewer-level data: 1) Review frequency, 2) Review timing and 3)Valence. Review frequency is measured by the number of reviews, which is used to identify segments referred to as "heavy" and "light" reviewers(Moon et al., 2010). Review timing is related to whether there exists a self-selection among reviewers in online product review(Li & Hitt, 2008). If reviewers at different period may have a different preference, such trend can affect the aggregated rating. Average rating of a reviewer reflects how favorable the viewer is when rating products in general. Specifically, this study examines how review frequency and review timing affect a reviewer's average ratings. In doing so, this study uses 42,172 posted online review messages generated by 6,941 identified reviewers for 59 movies released in the South Korea from July 2015 to December 2015. Review data were retrieved from NAVER movie review bulletin board, which serves as a popular portal site for movie goers. Moreover, NAVER provides review history data for each identified reviewer as shown in <Figure 1>, which enables us to collect review behavior data such as review frequency and timing.

<Figure 1> The Screen Shot of Review History for a Reviewer

This study contributes to the growing literature on the understanding how eWOM is generated at the individual consumer level. To best our knowledge, this is the first study that examine what reviewer's behavioral factors explain her/his tendency to vote positive or negative. On the basis of the main empirical findings, this study provides insights into building a recommendation system based on the consumer's review history data – frequency, timing, and valence. From a managerial perspective, our empirical findings suggest that firms should incorporate online reviewer's review frequency and timing when forecasting rating trend at the aggregate level and its potential sales impact. Beyond movie industry, our results can be applied to the other business area in which its market performance is significantly related to customer rating such as online retail stores.

The rest of the paper is organized as follows. The next section provides the discussions of the current literature and key research hypotheses. Then, the empirical model, data, results, and findings are discussed. Finally, the conclusion section includes the implications, limitations and future research.

2. Theory and Hypotheses

The vast literature on online review has focused on how online review affect market performance in the experience good industries such as movie (Elberse & Eliashberg, 2003; Liu, 2006; Dellarocas, Zhang, & Awad, 2007; Chintagunta, Gopinath, & Venkataraman, 2010), restaurant(Oluwafemi & Dastane, 2016), book(Godes & Mayzlin, 2004; Chevalier & Mayzlin, 2006), and mobile application(Lee, Wu, & Fan, 2017). Especially for the movie industry, movie goers actively involve interpersonal communication to increase accessibility of information and to influence others(Chaffee, 1982). Prior studies characterize the major two online reviews metrics -eWOM valence(user ratings) and eWOM volume(the number of reviews). Previous studies consistently document positive associations between eWOM volume and box-office sales, but the empirical evidence on the effect of eWOM valence on sales are somewhat contradictory(Godes & Mayzlin, 2004; Liu, 2006; Duan, Gu, & Whinston, 2008). For example, Godes and Mayzlin(2004) found a positive and significant effect of user ratings on sales for Amazon.com, but the insignificant effect for BarnsandNobles.com. Duan et al.(2008) pointed out that endogeneity between eWOM variables and market outcome leads to a mixed result for the eWOM valence.

Although vast of prior literature shed lights on the relationship between eWOM valence and sales outcome at the aggregate level, few studies have examined how individual viewer's reviewing characteristics influence on his/her average rating (valence). Online retailers try to collect individual-level rating for various products so that they can improve the accuracy of product recommendation system. Therefore, it is important to understand how a viewer's behavioral characteristics affect her/his overall attitude in movie rating. A previous study(e.g., Liu, Mai, & Yang, 2015) conducts a viewer-level analysis and document that a user's review writing experiences can affect negatively on the rating. Using individual viewer-level data, Moon et al.(2010) show that viewers' rating histories can affect viewer satisfaction. Numerical ratings for online customer reviews usually take interval scale which ranges from the lowest to the highest number set by the website managers. The lowest score indicates an extremely negative evaluation of the product while a very high rating reflects an extremely positive assessment of the product. At movie-level, Clemons et al.(2006) document the positive influence of favorable reviews on future sales. This study proposes that a viewer's review behavioral characteristics, namely, her/his review frequency and review timing can significantly affect the viewer's overall star rating. <Table 1> summarizes several studies document the individual reviewer-level characteristics.

<Table 1> Previous Studies on Online Review

Study	Data	Findings
Movie Level Analysis		
Elberse & Eliashberg (2003)	movie	Less positive reviews correspond to a higher number of opening screens, but more positive reviews mean more opening revenue
Godes & Mayzlin(2004)	book	A positive and significant effect of user ratings on sales for Amazon.com, but the insignificant effect for BarnsandNobles.com.
Chevalier & Mayzlin(2006)	book	Favorable reviews at one retailer lead to an increase in the sales of a product at the store relative to another retailer
Dellarocas, Zhang, & Awad(2007)	movie	First week WOM valence affect box office sales
Duan, Gu, & Whinston(2008)	movie	The influence of WOM volume on concurrent movie sales is positive. The influence of WOM volume on movie sales beyond the concurrent term is positive. However, the influence diminishes quickly.
Chintagunta, Gopinath, & Ventakraraman (2010)	movie	The main driver of box office performance is the volume of review
Reviewer Level Analysis		
Moon, Bergey, & Iacobucci(2010)	movie	Viewers' rating histories can affect viewer satisfaction
Liu, Mai, & Yang(2015)	movie	A user's review writing experiences can affect negatively on the rating.
This study (2017)	movie	A reviewer's review frequency and timing can affect his/her overall rating.

2.1. Review Frequency and Reviewer's Valence

A consumer's valence in product review may affect a product sale through affecting other potential consumers. Thus, it is meaningful to consider how a consumer's reviewing behavior affect his/her valence. Consumers can be segmented along a lot of dimensions based on their buying behavior. Prior studies have demonstrated that each market segment displayed different reactions to the marketing stimuli. Especially, the segment of heavy users is the core of any industry and are usually the heart of a successful marketing campaign. In this vein, previous research suggests that product experience is an important segmentation factor for on product attitude(Goolsbee & Klenow, 2002; Sorce, Perotti, & Widrick, 2005; Liu et al., 2015). Sorce et al.(2005) document that attitude in online buying is significantly different in frequency of buying goods. Liu et al.(2015) show that the effect of network externalities in online gaming reduces for the more experienced users who have participated in the game forum frequently. Similarly, frequent movie reviewers can gain knowledge by watching more movies.

We predict that frequent reviewers are less likely to have a favorable attitude movie scoring than light reviewers. First, heavy reviewers have more knowledge to appreciate the movie quality and able to utilize many features of a movie(i.e., genre, director, actors, plot, etc.) similar to movie critics. Moon et al.(2010) verify that viewers with more rating experience rate movies lower. As movie watching experience has accumulated, viewers can develop reliable judgment criteria and analyze movies similarly to professional critics. A prior study(i.e., Alba & Hutchinson, 1987) indicates that increased number of product-related experiences enhance the consumer's ability to evaluate the product-related features. Second, prior literature finds that expert rates a product based more on product-related aspects than on network externalities such as the number of viewers (Muthukrishnan & Wathieu, 2007; Liu et al., 2015). For hedonic product(i.e., theater, game, travel), more experienced viewers can appreciate the hedonic value of product characteristics than less experienced viewers(Nicolao, Irwin, & Goodman, 2009). Therefore, we hypothesize

<H 1> Heavy reviewers are less favorable in rating than light reviewers.

2.2. Review Timing on a Reviewer's Valence

A tendency to review early or late can influence how favorable attitude a viewer has toward movies on average. Prior study of the online word of mouth has shown inconsistent evidence for whether viewers are more likely to spread positive or negative information about products. Some findings suggest that viewers are more likely to engage in positive WOM, whereas others indicate that reviewers are more likely to engage in negative WOM depending on their review timing.

Prior study has argued that self-selection biases can arise in online review timing(i.e., Li & Hitt, 2008) because consumers hold difference preference. Early buyers and therefore early reviewers are more likely to have high valuations for the product quality compared to average buyers. Thus, the early movie reviewers may have strong intrinsic preferences for the directors, the actors, or the genre. Such early viewers may not have opportunities to correct their preference differences when interpreting few accumulate ratings and make purchase decisions while late reviewers may have enough time.

Valsesia, Nunes, and Ordanini(2016) suggest that being among the early reviewer can influence how an individual thinks about the product. They show that early reviewers are more positive due to a tendency to become psychologically attached to the product for which they can claim to be early adopters. When a viewer tends to leave a review early after movie release, s/he may have a high motivation to contribute the information transmitting process. Such findings are consistent with the self-enhancement hypothesis in online review literature. For example, Angelis et al.(2012) suggest that a consumer's self-enhancement motive leads her/his to generate positive WOM especially when they share information about their own positive consumption experience.

On the other hand, late viewers may hesitate to leave reviews online when they observe a significant number of previous reviews, because they may feel that they cannot contribute with new information about products. Supporting this, prior study Godes and Silva(2012) document that overall pattern of product rating is decreasing even after controlling for product and viewer effects. Furthermore, Moon et al.(2010) find that viewers' movie experience can cause them to become more critical in ratings over time. In sum, we expect that early reviewers tend to be positive in their ratings compared to late reviewers.

<H 2> Late period reviewers are less favorable than early period reviewers.

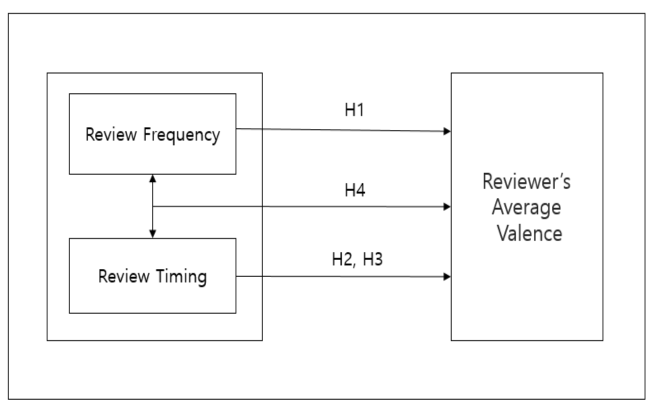
However, the very early reviewers who watch the movie right after release are more inclined to form higher expectations than other viewers. The initial viewers are only able to judge the quality of a movie based on their prior experience with the observable movie features such as genre, director, actor, and synopsis. Also, the movie distributors usually allocate their marketing budget right before release and during the opening week. Typically, a firm's the promotional message is biased to the positive part of the movie, it is more difficult for the first viewers to form a valid expectation on a new movie and leads to a higher expectation. As a result, the early viewers may provide lower ratings once a movie's actual quality does not meet their expectations. Hence, an inverted-U relationship can arise between a viewer's review timing and valence.

<H 3> (non-linear effect of review timing): A reviewer's average valence for movies increases with respect to the viewer's review timing and then decreases from a certain point.

Furthermore, there may be a more negative trend in the reviewer's average ratings due to the review frequency and review timing. Late reviewers participate the review after the quality of the movie is revealed by early and middle period reviewers. Thus, late reviewers form their expectation fully utilizing previously posted online review rating distribution, while early reviewers depend on their prior product experience or promotional messages by the firm. Once previous movie goers responses are available, the frequent reviewers who tend to write a review late could be the most skeptical group in movie rating. Hence, we hypothesize that the negative effect of review frequency should be more prominent for the late reviewers.

<H 4> The review frequency moderates the effect of review timing on a reviewer's valence. Heavy reviewers who tend to write reviews in the late period are likely to show lower rating.

<Figure 2> summarizes our model of reviewer's average valence. Frequent reviewers tend to give lower average ratings than occasional reviewers. Late reviewers are expected to provide less favorable ratings than early reviewers. Moreover, review timing has a non-linear relationship with average valence since the initial reviewers tend to show lower average ratings than middle period reviewers. Lastly, the interaction effect between review frequency and timing can arise because the heavy reviewers who leave a review late tend to show the lowest ratings on average.



<Figure 2> The Model of Reviewer's Average Valence

3. Empirical Applications

3.1. Data

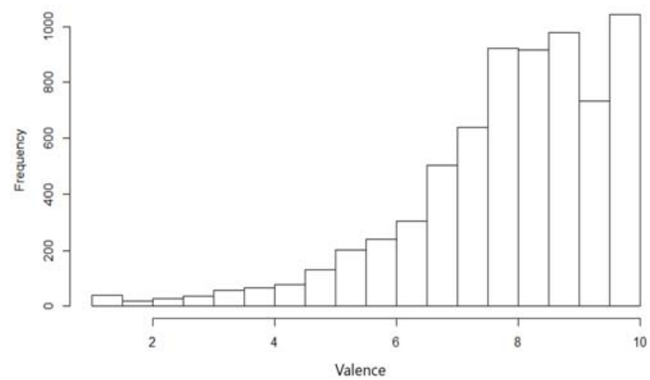
For this study, individual viewer-level data was collected through NAVER Movie(http://movie.naver.com). NAVER is the largest portal site to which more than 70% of population connects in South Korea on a daily basis. NAVER Movie also serves as the biggest movie portal site. It provides an identified viewer's review records including a rating(10 points scale system), text reviews no longer than 100 words, and review time stamp. For the 59 movies released from July 2015 to December 2015, 333,284 reviews written by 209,176 viewers were retrieved and parsed. Movie review bulletin board is still available to write after the movie is not shown in theaters. Therefore, the data includes reviews written only up to 8th week after release. Reviewers were identified based on their unique ID generating combination (e.g., 00(ABCD****)). Through this screening process, simple and duplicated viewer IDs were dropped.

<Table 2> illustrates the example of individual reviewer-level data. This study uses a hypothetical reviewer ID for illustration to disguise actual reviewer ID. This viewer had written 11 reviews out of 59 movies released during the sample period. The viewer-level data provides information on what rating the viewer has given for each movie(valence), how frequently the s/he writes a review(review frequency), and when s/he writes a review after a movie release(review timing). Although reviewer-level data cannot provide complete information on an individual viewer's movie watching behavior, our data can partially represent movie watching frequency and timing for a given period.

<Table 2> An Illustration of Review History Data for a Focal Viewer (Viewer ID = 00(ABCD****))

Movie	Country	Release Date	Review Date	Rating
Assassination	South Korea	2015-07-22	2015-08-06	8
Mission: Impossible - Rogue Nation	USA	2015-07-30	2015-08-28	7
O PISEU	South Korea	2015-09-03	2015-11-07	6
The Throne	South Korea	2015-09-16	2015-09-18	6
Maze Runner: Scorch Trials	USA	2015-09-16	2015-11-03	8
The Martian	UK/USA	2015-10-08	2015-10-28	8
The Advocate: A missing body	South Korea	2015-10-08	2015-11-14	7
The Phone	South Korea	2015-10-22	2015-12-08	6
Fatal Intuition	South Korea	2015-10-28	2015-11-09	7
The Priests	South Korea	2015-11-05	2015-12-21	7
The Hunger Games: Mockingjay-Part2	USA	2015-11-19	2015-12-04	8

To obtain the average valence of a viewer's historical movie ratings, this study restricted the sample into the reviewers who had written more than three reviews during the sample period. As a result, the final sample included a total of 42,172 reviews written by 6,941 reviewers for 59 movies. <Figure 3> illustrates the distribution of the reviewer's average valence. Due to self-selection bias, online review tends to show a J-shaped distribution with many high ratings and some low ratings(Moon, Park, & Kim, 2014). Our sample also shows a similar pattern. Out of total sample, the score greater than 9 is four times higher than the score less than 4. As a result, the mean score of 59 sample movies is relatively high(7.9 out of 10). The high average mean rating and J-shaped distribution represent that the viewers with high valuation tend to write reviews, while the viewers with low valuation tend not to write reviews. Thus, the valence distribution may not correctly represent the true evaluations of the total movie goers.



<Figure 3> Histogram of the Reviewer's Average Valence

The summary statistics for the variables in the reviewer-level dataset are shown in <Table 3>. The reviewers in the sample show a high level of average star rating ($VALENCE$) of 7.908. Review frequency($RFREQ$) represents the number of reviews written by a viewer during the sample period. On average, a reviewer in the sample leaves six reviews out of 59 movies during six months. Review timing ($RTIME$) stands for the average days after the release of a film when a viewer writes a review. On average, a viewer tends to write a review when 16 days have passed since the release date of a movie.

<Table 3> Summary Statistics (n=6,941)

Variable	Description	Mean	SD	Min	Max
$VALENCE$	Reviewer's Avg. Valence	7.908	1.678	1	10
$RFREQ$	Review Frequency	6.076	3.657	4	56
$RTIME$	Review Timing	16.05	10.454	1	62.5

3.2. Model

The movie rating system allows viewers to rate the quality of the film from the score from 1 to 10 as integer values. Due to the censored nature of movie score, this study adopts Tobit model specification to take into account the upper bound value of 10. Another reason for adopting Tobit model is to adjust the selection bias arising from the nature of J-shaped distribution of movie rating. When a selection bias exists, the probability of being included in the sample can be correlated with an explanatory variable (Heckman, 2013). In this case, OLS estimates can be biased. The censored nature of movie rating and the potential selection bias lead to a limited dependent variable. Thus, this study employed Tobit regression analysis to estimate the effect of a reviewer's behavioral characteristics on the average valence of the viewer. The goodness of fit is measured with the log likelihood.

The review frequency($RFREQ$) is expected to have a negative linear relationship with a reviewer's average valence for movies(<H1>). For the effect of review time($RTIME$), it is expected to have an inverted U-shaped relationship with a viewer's average star rating, implying that a viewer who tends to write a review during the middle period is likely to give more favorable responses than early or late period viewers(<H2>, <H3>). Thus both a linear term and a quadratic term of $RTIME$ need to be included. To test <H4>, an interaction term for review frequency and review timing is included. The interaction effect is expected to be negative because late viewers who frequently write reviews are likely to be skeptical in film rating. All explanatory variables take logarithm values to reduce the effect of extremes. As a result, the empirical model is as follows. $VALENCE_i$ indicates the mean rating of viewer i generated during the sample period.

$$VALANCE_i = \alpha + \beta_1 RFREQ_i + \beta_2 RTIME_i + \beta_3 RTIME_i^2 + \beta_4 RFREQ_i \times RTIME_i + \epsilon_i$$

4. Results

4.1. Tobit Estimation Results

<Table 4> shows the estimation results for nested models (M1~M5) of the average rating of a reviewer. To check robustness, an OLS estimation was also conducted and yielded similar results. That is, the ordinary regression model did not meaningfully affect the level of significance or the direction of the parameter estimates. Supporting <H1>, $RFREQ$ is negative and statistically significant($p < 0.000$) in M1 and M4. That is, review frequency affects negatively on reviewer's average rating. In addition, review time($RTIME$) has a significantly negative linear effect on the reviewer's average valence in M2. Thus, <H2> is supported. In M4 and M5, the positive effect of $RTIME$ and the negative effect of $RTIME^2$ indicates an inverted-U relationship between review timing and average rating. This implies that early and late period reviewers tend to show skeptical view than mainstream viewers who write reviews during the middle period. Therefore, <H3> is supported. The results of M5 also provide a strong support for <H3>, which hypothesizes that review timing moderates the effect of review frequency. The significant and negative interaction term, RFREQ×RTIME, indicates that frequent viewers who write a review during late the period are less favorable in film rating.

<Table 4> The Effect of Movie Reviewing Behavior on the Average Valence (n=6,941)

	Dependent variable: Average Valence				
	M1	M2	M3	M4	M5
$RFREQ$	-0.599***			-0.590***	0.269
	(0.055)			(0.055)	(0.182)
$RTIME$		-0.105***	0.520***	0.470***	1.007***
		(0.030)	(0.141)	(0.140)	(0.178)
$RTIME^2$			-0.134***	-0.122***	-0.113***
			(0.030)	(0.029)	(0.029)
$RFREQ \times RTIME$					-0.338***
					(0.068)
Constant	9.019***	8.265***	7.611***	8.666***	7.232***
	(0.096)	(0.080)	(0.164)	(0.190)	(0.347)
Log-Likelihood	-13347.27	-13401.47	-13391.48	-13332.91	-13320.76

Note: *p<0.1; **p<0.05; ***p<0.01

4.2. Predictive Value of Review Frequency and Timing

To test whether suggested model has a predictive value for estimating a reviewer's future valence, we merge the reviewer-level behavior information from the calibration set with the average valence of the corresponding reviewer in the validation set. The validation set consists of 19,330 reviews for 59 movies released from Jan 2016 to June 2016. They are written by 4,613 reviewers whose prior reviewing behavior information exits in the calibration sample(42,172 reviews written by 6,941 reviewers for 59 movies released from July 2015 to Dec 2015). Thus, the selected 4,613 reviewers have kept writing reviews in the next six months continuously. In <Table 5>, the correlations of review behavioral variables are high and significant. It implies that reviewers tend to show consistency in terms of their review frequency and timing. <Table 6> shows Tobit estimation results for the effect of movie reviewing behavior on the average valence in validation sample. The sign and significance of each independent variables are similar to those of calibration sample shown in <Table 4>. The estimation results suggest that the reviewer's prior behaviors can have predictive value for the reviewer's future average valence for a new movie. A summary of the hypotheses testing results are provided in <Table 7>

<Table 5> Correlation Test for Review Behavior Variables between Calibration and Validation Sample

Variables	Pearson	Spearman
$RFREQ_{cal} - RFREQ_{val}$	0.66***	0.42***
$RTIME_{cal} - RTIME_{val}$	0.51***	0.49***

<Table 6> The Effect of Movie Reviewing Behavior on the Average Valence (Validation set: n=4,613)

	Dependent variable: Average Valence				
	M1	M2	M3	M4	M5
$RFREQ$	-0.725***			-0.702***	0.019
	(0.086)			(0.086)	(0.289)
$RTIME$		-0.244***	0.465	0.377	0.834**
		(0.030)	(0.245)	(0.244)	(0.300)
$RTIME^2$			-0.151**	-0.130*	-0.119*
			(0.051)	(0.051)	(0.051)
$RFREQ \times RTIME$					-0.282**
					(0.108)
Constant	9.289***	8.625***	7.879***	9.199***	7.956***
	(0.159)	(0.080)	(0.286)	(0.328)	(0.577)
Log Likelihood	-9592.48	-9616.75	-9612.39	-9579.07	-9575.64
Note: *p<0.1; **p<0.05; ***p<0.01					

<Table 7> Summary of Findings

	Description	Variables	Parameter Estimates in <Table 4>	Result
H1	Heavy reviewers are less favorable in rating than light reviewers.	$RFREQ$	-0.590*** (M1)	Partially Supported
H2	Late period reviewers are less favorable than early period reviewers.	$RTIME$	-0.105*** (M2)	Supported
H3	(Inverted U-shaped Relationship) A reviewer's average valence for movies increases with respect to the viewer's review timing and then decreases from a certain point.	$RTIME$	1.007*** (M5)	Supported
		$RTIME^2$	-0.113*** (M5)	
H4	(Interaction Effect) The review frequency moderates the effect of review timing on a reviewer's valence. Heavy reviewers who tend to write reviews in the late period are likely to show lower rating.	$RFREQ \times RTIME$	-0.338*** (M5)	Supported

4.3. Comparison of Average Valence by Reviewer Group

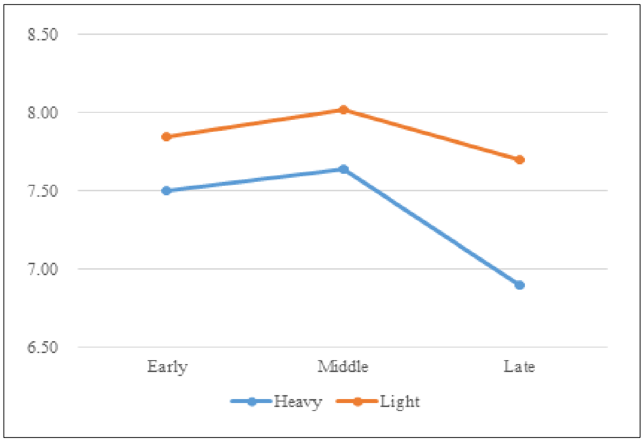
To illustrate the managerial implication in the film industry, the 6,941 viewers in the sample were categorized based on their review frequency and timing. Review frequency is measured as light and heavy depending on how many times they write reviews per month. We define a light reviewer who has written reviews less than or equal to once a month. Writing review more than once a month is considered as heavy reviewer's behavior in this study. As for the review timing, the reviewers who write reviews during opening week are categorized into early reviewers whereas the reviewers who write reviews after three weeks from the release are categorized into late reviewers. The reviewers who wrote the reviews between 2~3 weeks are considered middle period reviewers.

Descriptive statistics of average movie rating by the group are shown in <Table 7>. As shown in the table, the frequent reviewers tend to show lower average ratings than occasional reviewers. This shows that as the reviewers gain more experiences in watching movies and writing reviews, s/he tends to show more critical view. These results are consistent with prior studies(Li & Hitt, 2008; Moon et al., 2010). Consistent with Tobit model specification results, review timing has non-linear relationship with average movie ratings. That is, when a reviewer tends to write a review in the early period after release, s/he tend to show lower ratings than average reviewers. Then the middle period reviewers show a more generous view on movies and give higher ratings than average reviewers for the movie. However, the later reviewer who leaves reviews after three weeks after release tends to show skeptical views regardless of review frequency.

The analysis shows that among heavy movie reviewers, the rating tendency is significantly different depending on the timing of viewing. Late heavy movie reviewers show a more skeptical view on the movie(average rating=7.124) whereas early heavy movie reviewers show a less skeptical view on the movie(average rating=7.734). Both groups show skeptical view in average movie rating compared to sample average (=7.908). The difference of average rating between two groups(early heavy reviewers-late heavy reviewers) is 0.61 and statistically significant(t-stat=4.85, p<0.01). In addition, <Figure 4> shows the asymmetric impact of review timing depending on a viewer's review frequency.

<Table 7> Movie Reviewing Behavior and Average Valence(n=6,941)

Review Frequency	Review Timing	N	Mean	SD
Heavy	Early	438	7.734	1.623
	Middle	837	7.861	1.397
	Late	464	7.124	1.625
Light	Early	1254	7.864	1.874
	Middle	2681	8.134	1.568
	Late	1267	7.855	1.798



<Figure 4> The Effect of Review Frequency and Timing on Average Valence

5. Conclusions

5.1. Summary of Results

From managerial perspectives, this study shows that a reviewer's average valence is associated with his/her review frequency and timing. With individual reviewer-level analysis, this study highlights that frequent reviewers tend to show less favorable than occasional reviewers(<H1>). For review timings, later reviewers are likely to be negative than earlier reviewers on average(<H2>). However, inverted U-shaped relationship exists between review timing and average valence because reviewers during the second to third week from the movie release tend to show most favorable reviews(<H3>). Moreover, heavy reviewers who tend to write review in the later period are likely to show most critical view on movie ratings(<H4>).

5.2. Theoretical and Managerial Implications

This study contributes to both theory and practice. In theoretical perspectives, by building on self-selection theory, this study provides a conceptualization of what reviewer's characteristics explain the valence of the reviewer. We show that frequent reviewers are less favorable(<H1>). In addition, the late review groups show the negative(<H2>). Beyond movie industry, this study is particularly meaningful for online retailers whose sales outcome are highly dependent on the valence of online reviews. When a new product is introduced, the potential buyers with different preference may join writing reviews at different time period(<H3>). Among them, the heavy users of the product may have strict views on the new product and likely to provide an even lower review when they write a review at the later period(<H4>).

If viewers' tendency to review early and their likelihood of rating favorably is correlated, this self-selection behavior can cause higher rating in the early periods. In this study, we develop and empirically test a model that examines how review frequency and timing affect the idiosyncratic preferences of viewers. The hypotheses are tested using online movie reviews collected from NAVER movie. We find that on average, heavy reviewers tend to rate lower than light reviewers and late reviewers tend to rate lower than early reviewers. In addition, heavy reviewers who write reviews in the late period show the most skeptical rating than average reviewers.

In practice, the potential existence of self-selection bias suggests the unequal influence of the early buyers on market performance because the early reviews affect the quality perception of potential buyers and thereby affect the final market outcome(Li & Hitt, 2008). Therefore, online retailers need to take into account self-selection bias in eWOM generations and adjust marketing strategy accordingly. During the early period after a product launch or movie release, the reviewers with relatively high evaluation come and positively influence on ratings. Later on, the reviewers with more skeptical view on the new product come but their effect of following users might not be influential enough to affect potential sales. Understanding of individual reviewer-level behavior can help marketers to allocate resources for customer relationship management executions. This study suggests that early reviewers who write reviews less frequently might be the ideal target because they are likely to have positive perceptions.

5.2. Limitations and Future Research

There are some limitations of this research which can be investigated in the future research. First, this study only utilized summary ratings but didn't use the valence reflected in the textual reviews(Chevalier & Mayzlin, 2006). Future research could examine whether there exist self-selections in textual expressions with respect to review frequency and timing. Second, this study does not take into account the genre preference of individual viewers. Further investigation of moderating effect of genre on reviewer behavior could be an interesting research question. Third, the results for reviewer behavior may hold for the product like a movie which has experiential nature and new product adoption process. When individual reviewer-level data from online retail stores are available, our suggested model can be applied to test whether product category moderates the effect of review frequency and timing on the reviewer's valence. For example, do heavy reviewers in search goods tend to show more negative valence than those in the experiential goods? Do late reviewers in search goods tend to show more negative valence than those in experiential goods? While viewing a movie involves more emotional consumption, purchasing a search good such as a television involves more rational consumption. Therefore, comparing differences in review behavior for the two types of goods might be a meaningful research work. We leave these questions for future research.

References

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of consumer research*, *13*(4), 411-454.

Angelis, M. D., Bonezzi, A., Peluso, A. M., Rucker, D. D., & Costabile, M. (2012). On braggarts and gossips: A self-enhancement account of word-of-mouth generation and transmission. *Journal of Marketing Research*, *49*(4), 551-563.

Bayus, B. L. (1985). Word of mouth-the indirect effects of marketing efforts. *Journal of advertising research*, *25*(3), 31-39.

Chaffee, S. H. (1982). Mass media and interpersonal channels: Competitive, convergent, or complementary. *Inter/media: Interpersonal communication in a media world*, *57*, 77.

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

Chintagunta, P. K., 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 Science*, *29*(5), 944-957.

Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems*, *23*(2), 149-171.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing*, *21*(4), 23-45.

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter?—An empirical investigation of panel data. *Decision support systems*, *45*(4), 1007-1016.

Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, *22*(3), 329-354.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291-313.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing science*, *23*(4), 545-560.

Godes, D., & Silva, J. C. (2012). Sequential and temporal dynamics of online opinion. *Marketing Science*, *31*(3), 448-473.

Goolsbee, A., & Klenow, P. J. (2002). Evidence on learning and network externalities in the diffusion of home computers. *The Journal of Law and Economics*, *45*(2), 317-343.

Heckman, J. (2013). Sample selection bias as a specification error. *Applied Econometrics, 31*(3), 129-137.

Hu, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, *9*(3), 201-214.

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

Lee, J. H., Wu, R., & Fan, L. (2017). Influence on the use intention of user's traits in China market. *East Asian Journal of Business Management*, *7*(2), 21-29.

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456-474.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, *70*(3), 74-89.

Liu, Y., Mai, E. S., & Yang, J. (2015). Network externalities in online video games: An empirical analysis utilizing online product ratings. *Marketing Letters*, *26*(4), 679-690.

Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic effects among movie ratings, movie revenues, and viewer satisfaction. *Journal of Marketing*, *74*(1), 108-121.

Moon, S., Park, Y., & Seog, Kim, Y. (2014). The impact of text product reviews on sales. *European Journal of Marketing*, *48*(11/12), 2176-2197.

Muthukrishnan, A. V., & Wathieu, L. (2006). Superfluous choices and the persistence of preference. *Journal of Consumer Research*, *33*(4), 454-460.

Nicolao, L., Irwin, J. R., & Goodman, J. K. (2009). Happiness for sale: Do experiential purchases make consumers happier than material purchases?. *Journal of Consumer Research*, *36*(2), 188-198.

Neelamegham, R., & Jain, D. (1999). Consumer choice process for experience goods: An econometric model and analysis. *Journal of marketing research*, 373-386.

Oluwafemi, A. S., & Dastane, O. (2016). The Impact of Word of Mouth on Customer Perceived Value for the Malaysian Restaurant Industry. *East Asian*

Journal of Business Management, *6*(3), 21-31.

Smith, D., Menon, S., & Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of interactive marketing*, *19*(3), 15-37.

Sorce, P., Perotti, V., & Widrick, S. (2005). Attitude and age differences in online buying. *International Journal of Retail & Distribution Management*, *33*(2), 122-132.

Sthapit, A., Jo, G. Y., & Hwang, Y. Y. (2016). Construal Levels and Online Shopping: Antecedents of Visits to and Purchases from Online Retailers' Websites. *Journal of Industrial Distribution & Business*, *7*(3), 19-25.

Valsesia, F., Nunes, J., & Ordanini, A. (2016). I Got Here First! Feelings of Psychological Ownership and Consumer Ratings. *ACR North American Advances,* Retrieved May 21, 2017 from https://thescienceofownership.org/category/general/