

Association for Information Systems AIS Electronic Library (AISeL)

AMCIS 2009 Doctoral Consortium

Americas Conference on Information Systems
(AMCIS)

2009

Competing for Attention in Online Reviews

Wenqi Shen

Purdue University, shenw@purdue.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2009_dc

Recommended Citation

Shen, Wenqi, "Competing for Attention in Online Reviews" (2009). *AMCIS 2009 Doctoral Consortium*. 2.
http://aisel.aisnet.org/amcis2009_dc/2

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2009 Doctoral Consortium by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Competing for Attention in Online Reviews

Wenqi Shen

Ph.D. Candidate

Krannert School of Management

Purdue University

West Lafayette, IN 47907-2056

shenw@purdue.edu

ABSTRACT

Millions of web users engage in the online activities such as blogging, online forums, or online review systems to interact with people and to capture attention. This study tries to understand how online users compete for the scarce resource, attention, when participating in the online Web 2.0 activities. We develop a framework to capture the decision making process for online users to choose a right topic to post information and right content to post. Using book reviews from Amazon, we find that online reviewers do behave rationally in order to gain attention and to enhance their reputation. Our results suggest that experienced or top ranking reviewers are more likely to review relatively obscure books to avoid severe competition for attention in popular books. Moreover, top ranking reviewers usually post reviews earlier than low ranking reviewers as there are fewer reviews coexisting at the early stage to compete for attention. In terms of review ratings, we find that low ranking reviewers post more extreme ratings which distance themselves from the current average rating.

Keywords

Attention, online reviews, Web 2.0, virtual community, user-generated content.

INTRODUCTION

In the information age, there arises a new kind of economy, namely the attention economy. Millions of web users are engaging in the Web 2.0 activities such as blogging, online forums, or online review systems to offer information in digital forms so as to interact with people and capture attention. These online activities are typically voluntary activities without any direct monetary rewards. However, even with no monetary incentives we still observe millions of users making efforts to participate in these activities. For example, there are over 3 million online reviewers registered on Amazon.com to offer product reviews. Studies have shown that online users are typically driven by social capital such as peer recognition (i.e. peer attention), or online reputation, to keep participating in the online activities voluntarily (e.g. Jeppesen and Frederiksen 2006, Lerner and Tirole 2002).

While we are aware that the underlying motivation for online users' contributions is to gain attention and reputation, researchers pointed out that this increasing digital information creates the processing problem for users to seek relevant and useful information (Hansen and Haas 2001, Reuters 1998, Hunt and Newman 1997). With the large supply of user generated content online, the scarce resource is not the information itself, but the limited attention that online users are able to allocate on each of the online posts (Ocasio 1997). For example, Simon (1997) argued that "a wealth of information creates a poverty of attention". Given the scarce nature of the attention and the incentive of gaining it, online users are likely to compete among one another to gain attention when they post information online.

In this paper, we focus on understanding how online users compete for attention when they post information. In particular, we try to address the following research questions. How do online users strategically compete against one another to obtain more attention? Do they strategically choose certain topics to participate in, for example popular topics or unpopular topics? When posting information, do they choose a safe strategy by following mass opinions or do they choose a risky strategy by differentiating their opinions?

We use online reviewers' behavior as the context to study these research questions. Online reviews play an important role in consumers' purchasing decisions (e.g. Chevalier and Mayzlin 2006, Liu 2006). This suggests that not only are millions of online reviewers participating in the reviewing activities, but also that a large amount of potential consumers or online users

are seeking review information for their purchasing decisions. According to a survey conducted by Deloitte's Consumer Products Group in 2007, 62% of consumers read online product reviews, and 80% of them say reviews influence their purchase decisions. Since peers' attention and online reputation motivates online review activity, online reviewers will have to strategically provide valuable reviews so as to compete for these scarce resources. Therefore, online review activity suits the purpose of this study.

Using book reviews from Amazon, we find that online reviewers do behave rationally in order to gain attention and to enhance their reputation. Our dataset contains all fiction books released in October and November 2008. One unique feature of the dataset is that we collect daily information for each book from its release date so that we are able to identify reviewers' behaviors throughout the time frame of the study. We find that top ranking reviewers are usually more strategic players than low ranking reviewers in online review activity. In terms of which topic (book) to participate in, our results suggest that top ranking reviewers are more likely to review relatively obscure books to avoid severe competition for attention in popular books than low ranking reviewers. Moreover, top ranking reviewers usually post their reviews earlier than low ranking reviewers as there are fewer reviews coexisting at the early stage to compete for attention. In terms of review ratings, we find that low ranking reviewers are more likely to post extreme ratings which differentiate themselves from the current average rating as the cost of hurting reputation is relatively low. This is probably the strategy adopted by low ranking reviewers to distinguish themselves so as to build up their online reputation.

The rest of the proposal is organized as follows. We first summarize the previous literature and identify the gap in the existing literature that this study tries to fill. Next, we develop a conceptual framework to model the decision process for online users to participate in the online activities so as to gain attention. Then, we present the data and the methodology in section 4 and 5. Finally, we show the preliminary results followed by a discussion of the implications in the conclusion section.

LITERATURE REVIEW

With the growing popularity of Web 2.0 activities and the business values generated from these activities, there is more interest among academia in studying these online activities and their impacts (e.g. Aggarwal et al. 2006, Aggarwal et al. 2007, Scoble and Israel 2006, Mayzlin 2006). For example, there is a large body of literature addressing the impacts of online reviews on product sales (e.g. Basuroy et al. 2003, Dellarocas et al. 2004, Li and Hitt 2008, Zhang et al. 2004). They mainly use two measures to study the impacts of online reviews on product sales, *Volume* and *Valence* (e.g. Liu 2006, Zhang et al. 2004).

Volume measures the number of online reviews (e.g. Chevalier and Mayzlin 2006). A high volume of online reviews can increase the awareness of a product among potential buyers and therefore can increase product sales (Liu 2006). Valence measures the positive or negative nature of online reviews. Unlike volume, the impact from the valence of online reviews is mixed. For example, using user reviews on Yahoo! Movies, Liu (2006) and Duan et al. (2005) found that the valence of previous movie reviews does not have significant impact on later weekly box office revenues. However, Zhang and Dellarocas (2006) found a significant positive relationship between the valence of online reviews and box office revenues.

Recently, researchers have started to consider the impact of web users' characteristics in addition to the numerical aspects of the reviews on product sales. For example, Forman et al. (2008) considers the effect of users' online identities on the impact of reviews. They report that reviews posted by real name users will have a larger impact on product sales than those by anonymous users. Chen et al. (2006) uses reviewers' ranking to measure the impact from reviewers' online reputation on the impact of their reviews. They find that reviews by top ranking reviewers will have a more powerful impact on product sales than those by low ranking reviewers.

While understanding the consequences or the impacts of these online posts is very important, it is also important to know how online users behave when they participate in these online activities, especially without any monetary rewards. In the literature of member contributions in online communities, studies have found that when lacking monetary incentives, social incentives such as peer recognition and online reputation are important drivers for community members to contribute voluntarily (e.g. Lerner and Tirole 2002). For example, in the open source software context, Lerner and Tirole (2002) found that reputation and peer recognition are the primary motivations for the providers to contribute to the community without monetary rewards. They claimed that the main driver of providers' efforts is the "reputation capital" they gained by contributing to the community. Providers' reputation signals their competence which drives them to participate online. In the context of firm-hosted user communities, firm recognition of user contributions is also reported as valuable to the users

(Jeppesen and Frederiksen 2006). Positive reputation and peer recognition can motivate participants to keep contributing voluntarily (Pavlou and Gefen 2004, Resnick et al. 2000).

These theories and findings are applicable to the Web 2.0 communities where users' contributions are typically motivated by peers' attention and their online reputation. For example, in an online review system, reviewers have to devote substantial amount of time and efforts to write reviews. However, they typically do not get any monetary rewards for their contributions. This type of community is similar to the open source software development community as mentioned above. Based on the findings in the open source software literature, we argue that reviewers' reputation in terms of their rankings would be viewed as an important reward to the reviewers. In other words, they post reviews to attract more attention and gain reputation to reward their efforts.

While we are aware of the impacts of online posts and the motivations for online users to generate these posts, little has been known on how these social incentives and motivations can affect users' actions and decisions. In other words, there is a gap in the literature to link the underlying motivations for users to contribute and the outcome of their contributions, i.e. the online posts they generate. These outcomes can significantly affect the consequent impact of these online activities such as the impact of online reviews on product sales. In this study, we try to fill in this gap by developing a framework to study the decision making process for web users when participating in online activities to gain social benefits.

CONCEPTUAL FRAMEWORK AND HYPOTHESES

The conceptual framework developed in this study tries to capture the decision making process of web users, online reviewers in particular, when they join various post-based Web 2.0 activities such as online reviews. When reviewers try to post a review to gain attention, they are facing two sequential decisions: first, which topic or product to choose, and second, what content to post.

Topic Choice

The topic choice problem is for the reviewers to choose an ideal topic that can maximize potential attention. This involves balancing between the popularity of the topic and the level of competition for attention within that topic. A popular topic usually can attract much attention from online users. For example, the top seller books on Amazon can generate more than one thousand reviews while an obscure book may or may not have any reviews. In other words, the level of total attention for a popular topic is usually much higher than that for an unpopular topic.

However, choosing a popular topic can bring more severe competition for attention at the same time. Since there are more posts competing for attention within the popular topic, it is not easy for one individual post to outperform other posts to gain adequate attention. One can easily drown in the flood of similar posts. In other words, although the level of total attention for that topic is high, the individual share of attention could be low. This effect would be more significant if attention is highly concentrated. However, if reviewers choose to participate in some unpopular topics, they may easily become a big fish in a small pool. Therefore, the obscure topics may seem unattractive at the first place. They may turn out to be rewarding in terms of potential individual attention to the users.

Considering both effects from the popularity of topics and the competition for attention, experienced reviewers will try to avoid contributing to highly popular topics which already have many posts. They will be more likely putting efforts to those topics that are either less popular or have only a few existing posts. However, inexperienced web users may not be able to construct such sophisticated strategy when choosing topics. Often time, they will be following the hot topics without considering too much about how to maximize the potential attention their posts can gain. Therefore, we expect the following hypotheses.

H1: Experienced reviewers will choose less popular books to review than inexperienced reviewers will do.

H2: Experienced reviewers will post reviews when fewer reviews exist than when inexperienced reviewers will do.

Content Choice

After choosing a topic, reviewers have to find a way to compete with other posts through their content. On one hand, offering a unique opinion to differentiate from the rest of the posts could be an effective way to capture attention. For example, it could be eye-catching to post an extreme rating that is far away from the average rating on a review site or to use some special phrases that have not been used before. Therefore, using a differentiation strategy may help the post to stand

out. However, this could be a risky strategy that may lead to disagreement and bring negative feedback. For example, Postmes et al. (2000) pointed out that online community members are desire for behaviors that are consistent with the community norms. Reviewers' posts that are patterned after community norms are easy to communicate to others and establish their reputation (Forman et al. 2008, Resnick et al. 2000). Therefore, using a differentiation strategy might hurt their reputation and discourage online users.

On the other hand, following mass opinions or imitating others' posts may seem to be a safe strategy to certain users. Although such a post may not be able to draw as much attention as the unique one, it will not bring much negative feedback at least. For users whose cost of hurting reputation is high, an imitation strategy may be more attractive than a differentiation strategy. Moreover, if the current competition for attention is relatively low, a safe strategy, i.e. the imitation strategy, may result in high expected returns.

Online users will face the question of which strategy to choose to gain more attention. Reviewers with established reputation may not be willing to offer extreme opinions as the cost of hurting reputation is relatively high. Novices may have to take the risk of providing unique opinions to distinguish themselves so as to capture attention and establish their online reputation. In addition, reviewers may not have to choose the risky strategy when posting a review early as there is limited competition for attention at early stage. However, if they post reviews late, they will have to take the risk by differentiating their posts since the competition for attention becomes high. Therefore, we hypothesize reviewers' content choice as follows:

H3: Reviewers with high reputation will be more likely to use the safe strategy by posting reviews that are closer to the average reviews than reviewers with low reputation.

H4: Reviewers will be more likely to use the safe strategy by posting reviews that are closer to the average reviews when they review a book early than when they review a book late.

Figure 1 summarizes the conceptual framework which models the two-step decision making process that reviewers will take when posting a review to gain attention.

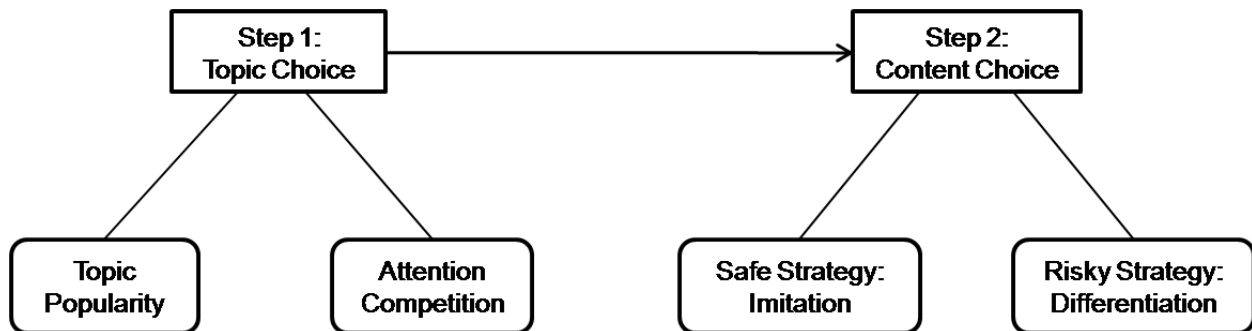


Figure 1. Conceptual Framework

DATA

This study uses book reviews on Amazon.com. We select Amazon as it is the leading electronic retailer for books which represents 70% of the whole market transactions. It has also been chosen to study research questions regarding online reviews by various previous researches (e.g. Chen et al. 2006, Forman et al. 2008, Li and Hitt 2008). Our sample includes all fiction titles released between October 1 and November 31 2008 which contains about 1400 books. We choose fiction as it is one of the top book categories on Amazon which usually attract adequate reviews for our analysis. In addition, fiction is also among the categories which have a relatively large amount of new releases every month.

The data in our sample includes daily information on books, reviews, and reviewers. For books, we collect the book's daily price and sales rank which will be used as a proxy of its actual sales volume. For reviews, we collect the date when the review is posted, the reviewer's user name which could be a real name or a pen name, the review rating, the helpful vote (this indicates how many readers find this review helpful) and the total vote (this is used as a proxy of the amount of attention it has captured). The votes are collected daily. Based on the reviews, we then obtain the information from each reviewer's online profile on Amazon. This includes the reviewers' user name, the total number of reviews they have posted in history, and their reviewer rank on Amazon (Amazon ranks reviewers according to the number of reviews and the quality of their

reviews¹). Again, we track the daily changes of reviewers' profiles such as their ranks, the total number of reviews they have posted, and the total helpful votes they receive. The following table summarizes the data in our sample. One unique feature of our sample is that we collect all the information from the release date of the books. Therefore, we are able to observe reviewers' strategies along the time. The data expands a three-month period from October 1 2008 to January 1 2009.

METHODOLOGY

Topic Choice

We model user's topic choice by using the multinomial logit model which has been applied extensively in the marketing literature to study consumers' brand choice (e.g. Carpenter and Lehmann 1985, Gupta 1988, Bell et al. 1999). It is suitable for studying the topic choice since it is based on a behavioral theory of utility, accounts for explanatory variables, and allows competition (Gupta 1988). We define the topic choice model as follows:

$$P_{rkt} = \frac{\exp(b'X_{rkt})}{\sum_{i=1}^B \exp(b'X_{rit})}, \quad k = 1, \dots, B \text{ books}$$

where:

P_{rkt} = probability that reviewer r chooses book k to review on day t and

X_{rkt} = the expected attention reviewer r can get for reviewing book k on day t and

$$X_{rkt} = c'Y_{rkt} + \varepsilon$$

where:

Y_{rkt} = the vector of explanatory variables that may affect the potential attention reviewer r can get for reviewing book k on day t .

The explanatory variables, Y_{rkt} , include the log transform of sales rank ($\ln SalesRank$), number of existing reviews ($\ln ReviewNum$), the number of days from the release date ($\ln DaysElapsed$), reviewer rank ($\ln ReviewerRank$), and average current rating ($Rating$). We will first estimate the expected attention a reviewer can get and then estimate the logit model to understand what factors affects reviewer's choice of topics, i.e. books. To test H1 and H2, we will be interested in the coefficient of $\ln SalesRank$ and $\ln ReviewNum$.

Content Choice

After selecting a topic, users will consider what content to post. For example, in the review context, the reviewers will decide the rating and the review content. They can either choose a safe strategy, which is to imitate previous posts, or a risky strategy, which is to differentiate from previous posts. In other words, for the ratings, reviewers can either post a rating which is close to the current average rating or an extreme rating which is far away from the average rating. For the review content, they can either repeat the product features and the attitude towards these features as in the previous reviews or they can describe some new features or use a different tone. Next, we describe the variables used to study reviewers' content choices.

Dependent Variables

To measure whether reviewers imitate or differentiate when posting ratings, we define a dependent variable, *Valence*, which is the difference between one reviewer's rating and the current average rating. *Valence* captures the distance of that reviewer's rating and the previous average rating. A small value of *Valence* suggests that the reviewer is imitating, while a large value indicates differentiation. In addition, we use $Valence^2$ to measure the absolute distance between the reviewer's rating and the average rating which ignores the positive or negative nature of the difference.

$$Valence = Rating - Average Rating$$

$$Valence^2 = (Rating - Average Rating)^2$$

Independent Variables

Since our objective is to study how different reviewers compete for attention, we use two variables to control for reviewer characteristics, the reviewer rank, $\ln(ReviewerRank)$, and the real name identity, *RealName*, for each reviewer. *RealName* is a dummy variable which takes 1 if the reviewer uses a real name and 0 otherwise. When the reviewer posts a review on

¹ Amazon uses the ratio of Helpful vote/Total vote to measure the quality of a review. In addition, they claimed that they considered the relative magnitude of the amount of total vote at the same time.

Amazon, Amazon will also display a real name badge for the reviewer who uses a real world name and a top reviewer badge if the reviewer is in the top 1000, 500, 100, 10, or 1 reviewer ranking list. Therefore, these two variables capture the basic indicators that can distinguish among reviewers. Moreover, using reviewer rank can offer insights on how social incentive such as online reputation affects users' decisions.

Meanwhile, to identify whether a reviewer imitate previous posts, we use the number of days after the release date of the book, $\ln(\text{DaysElapsed})$, to measure the effect of the time. We control for the popularity of the book as reviewers may also behave differently for popular books versus obscure books. We use the sales rank of the book, $\ln(\text{SalesRank})$, as a proxy of the popularity of the books.

The models we will run to test H3 and H4 are as follows:

$$\text{Valence}_{kr} = \beta_0 + \beta_1 \ln(\text{ReviewerRank}_{kr}) + \beta_2 \text{RealName}_{kr} + \beta_3 \ln(\text{DaysElapsed}_{kr}) + \beta_4 \ln(\text{SalesRank}_k) + \mu_k + \varepsilon_{kr}$$

$$\text{Valence}^2_{kr} = \beta_0 + \beta_1 \ln(\text{ReviewerRank}_{kr}) + \beta_2 \text{RealName}_{kr} + \beta_3 \ln(\text{DaysElapsed}_{kr}) + \beta_4 \ln(\text{SalesRank}_k) + \mu_k + \varepsilon_{kr}$$

where μ_k is a book fixed effect which controls for the differences in the rating valence across books. As stated in H3, we expect the coefficient for $\ln(\text{ReviewerRank})$, β_1 , to be positive and significant. For H4, we expect the coefficient of $\ln(\text{DaysElapsed})$, β_3 , to be positive and significant.

PRELIMINARY RESULTS

We present our preliminary results in this section. Table 1 summarizes the descriptive statistics for the variables. There are 338 books that have at least one review by the end of the data collection period and 2926 reviews in total.

Variable	Obs.	Mean	Std. Dev.	Min	Max
<i>Number of Reviews</i>	2926	44.82	58.857	1	290
<i>lnReviewNum</i>	2926	2.89	1.486	.69	5.67
<i>SalesRank</i>	2915	69850.49	256201.700	3	7423646
<i>lnSalesRank</i>	2915	7.05	3.432	1.09	15.82
<i>lnDaysElapsed</i>	2926	3.18	.994	0	4.71
<i>RealName</i>	2926	.58	.494	0	1
<i>Rating</i>	2926	3.98	.780	1	5
<i>lnReviewerRank</i>	2065	11.45	3.324	.69	15.47
<i>Valence</i>	2926	-.06	1.120	-3.73	3.04
<i>Valence</i> ²	2926	1.26	2.045	0	13.98

Table 1. Descriptive Statistics

In terms of which topic (book) to choose, we use *t*-test to compare whether experienced or top ranking reviewers pick books differently than low ranking reviewers. Table 2 suggests that top ranking reviewers are more likely to review relatively obscure books to avoid severe competition for attention in popular books. Moreover, we also compare whether top ranking reviewers intend to pick a certain time to post the reviews. Table 3 and 4 show that top ranking reviewers usually post a review earlier than low ranking reviewers as there are fewer reviews coexisting at the early stage to compete for attention. These results suggest that top ranking reviewers will strategically choose an appropriate book and a good time to review so as to gain more attention. These preliminary findings are consistent with the prediction of our framework and the hypotheses, H1 and H2, that reviewers try to balance between the popularity of the book and the competition for attention when choosing books. Based on these findings, we will conduct more analyses using the topic choice model developed in the methodology section to obtain further insights.

Reviewer Rank	Top 100	Low ^a	Top 500	Low	Top 1000	Low
Mean	8.92	6.99	8.51	6.98	8.27	6.96
Std. Dev.	3.407	3.417	3.387	3.419	3.452	3.414
Obs.	88	2827	132	2783	193	2722
Diff (Top – Low)	1.93**		1.53**		1.31**	

Note: ** p -value < 0.01, * p -value < 0.05

a. The Low group contains the rest of the reviewers whose reviewer rank is larger than the rank in the Top group.

Table 2. t -test for mean $\ln(\text{SalesRank})$

Reviewer Rank	Top 100	Low	Top 500	Low	Top 1000	Low
Mean	1.37	2.94	1.68	2.95	1.87	2.96
Std. Dev.	1.033	1.472	1.226	1.472	1.241	1.475
Obs.	89	2837	133	2793	194	2732
Diff (Top – Low)	-1.57**		-1.27**		-1.09**	

Note: ** p -value < 0.01, * p -value < 0.05

Table 3. t -test for mean $\ln(\text{ExistingReviewNumber})$

Reviewer Rank	Top 100	Low	Top 500	Low	Top 1000	Low
Mean	1.84	3.23	2.23	3.23	2.42	3.24
Std. Dev.	1.107	.961	1.265	.957	1.198	.956
Obs.	89	2837	133	2793	194	2732
Diff (Top – Low)	-1.38**		-1.00**		-.82**	

Note: ** p -value < 0.01, * p -value < 0.05

Table 4. t -test for mean $\ln(\text{DaysElapsed})$

In terms of review content, we find that low ranking reviewers tend to post more extreme ratings which distance themselves from the current average rating, which is consistent with H3. Table 5 shows that the coefficients for $\ln(\text{ReviewerRank})$ are positive and significant in Model 2 and 4². This supports the prediction in H3 that low ranking reviewers tend to use the differentiation strategy to distinguish themselves so as to capture more attention, while top ranking reviewers tend to use a safe strategy by not posting significantly deviated ratings.

The coefficients for $\ln(\text{DaysElapsed})$ are positive and significant in Model 2 and 4. These results are consistent with H4 that when reviewers post reviews late they will post more extreme reviews than when they post reviews early.

Comparing Model 3 and 4 with Model 5 and 6, we find that top ranking reviewers behave differently from low ranking reviewers. Low ranking reviewers are more likely to use a differentiation strategy but it is not obvious among top ranking reviewers. For example, for top reviewers, they seem not to post extreme reviews even when they post reviews late, i.e. the coefficient for $\ln(\text{DaysElapsed})$ is not significant in Model 5 and 6. This may be due to the fact that top ranking reviewers typically post reviews earlier than low ranking reviewers as discussed above.

These findings are consistent with the prediction from our framework that top reviewers tend to be more conservative than the low ranking reviewers as the cost of hurting reputation is much higher to them. Therefore, top ranking reviewers will be less likely to choose a risky strategy, i.e. the differentiation strategy.

² The model fit for Model 1 and Model 3 is extremely low, so we do not interpret the results from these two models.

	Model 1 ^{a,i}	Model 2 ^{b,i}	Model 3 ^{a,ii}	Model 4 ^{b,ii}	Model 5 ^{a,iii}	Model 6 ^{b,iii}
<i>ln(ReviewerRank)</i>	-.02* (.009)	.07** (.016)	-.02 (.014)	.08** (.024)	-.04 (.038)	-.02 (.052)
<i>RealName</i>	.11* (.056)	-.07 (.097)	.11 (.061)	-.08 (.106)	.11 (.191)	.25 (.260)
<i>ln(SalesRank)</i>	.03 (.043)	-.03 (.075)	.02 (.048)	-.03 (.084)	-.04 (.149)	-.11 (.203)
<i>ln(DaysElapsed)</i>	.02 (.038)	.18** (.066)	-.02* (.043)	.19** (.076)	.17 (.102)	.26 (.139)
Obs.	2057	2057	1864	1864	193	193
R²	.06	.18	.06	.17	.35	.36

Note: ** p -value < 0.01, * p -value < 0.05

a. Model 1, 3, and 5 use *Valence* as the dependent variable.

b. Model 2, 4, and 6 use *Valence*² as the dependent variable.

i. Model 1 and 2 contain the whole data set.

ii. Model 3 and 4 contain reviewers whose rank is larger than 1000, i.e. the low ranking reviewers.

iii. Model 5 and 6 contain reviewers whose rank is above 1000, i.e. the top ranking reviewers.

Table 5. *Valence* and *Valence*²

CONCLUSIONS

Using book reviews from Amazon, we show reviewers with different reputation level exhibit different preference on book choices and content choices. Reviewers with more experience and high reputation level are more rational users whose decisions are consistent with the prediction of our framework. Reviewers with less experience and low reputation level are not as rational as top ranking reviewers when choosing books to review. Moreover, low ranking reviewers are more willing to take the risky strategy than top ranking reviewers by offering extreme ratings.

To the best of our knowledge, this study is the first attempt to understand how online users' decisions are driven by the rational desires to gain attention and online reputation. It fills the gap in the literature on how the social motivations can actually affect online users' decisions. This study offers several insights to companies who try to create social incentives to encourage participation from web users.

First, our results indicate that web users are serious about these social benefits and their behaviors could be influenced by these pre-designed mechanisms such as reviewer ranking system. Adding such social benefits to the users may actually encourage more serious usage of the online system and more contributions from the web users.

Second, companies may utilize these incentives to manipulate users' behaviors to achieve their business goals. For example, since reviews posted by top reviewers are usually more powerful, if companies want to promote certain types of products, they can offer more social incentives to entice reviewers to review these products. For instance, they can double the weights on these products when computing the reviewer rank. In that case, reviewers who treasure their online reputation, typically top ranking reviewers, will review these products more actively. As a result, these products can gain more attention from potential buyers.

Third, not only are the topic choice affected by the social incentives, the content they offer are also affected. This, to some extent, offers opportunities for companies to predict or control for user generated content. For example, we find that top ranking reviewers are less likely to offer extreme ratings. If companies can offer users enough social incentives which add adequate cost for them to post a deviated rating or extremely negative rating, these unwanted negative ratings can be reduced. In other words, angry customers may be more tolerant to the unsatisfied products as they do not want to destroy their build-up reputation.

The future direction of this study includes exploring the topic choice model to formally test H1 and H2 and the text content of the reviews. Not only do review ratings reveal the strategic choice that reviewers are making, but also the text content can affect the level of attention that reviews can gain. Therefore, reviewers will need to determine a strategy for writing the content as well. For example, they can repeat the product features and the opinions that early reviews have discussed or raise

some new issues. We develop a measurement in Appendix A to measure the frequency of commonly used words of each review to study reviewers' decisions on review content. Combining the results from the ratings and the text content will bring us a holistic picture of reviewers' decisions when offering reviews.

REFERENCES

1. Aggarwal, R., Gopal, R. and Sankaranarayanan, R. (2006) Blog, Blogger and the Firm: An Analysis of Employee Incentives and Firm Policies, in *16th Workshop on Information Technologies and Systems (WITS)*.
2. Aggarwal, R., Gopal, R. and Gupta, A. (2007) Putting Your Money Where Mouth is: Role of Electronic Word-of-Mouth on New Venture Funding, in *17th Workshop on Information Technologies and Systems (WITS)*.
3. Basuroy, S. and Chatterjee, S. and Ravid, S.A. (2003) How Critical are Critical Reviews? The Box Office Effects of Film Critics Star-Power, and Budgets, *Journal of Marketing*, 67, 4, 103-117.
4. Bell, D.R., Chiang J. and Padmanabhan, V. (1999) The Decomposition of Promotional Response: An Empirical Generalization, *Marketing Science*, 18, 4, 504-526.
5. Carpenter, G.S. and Lehmann, D.R. (1985) A Model of Marketing Mix, Brand Switching, and Competition, *Journal of Marketing Research*, 22, 318-329.
6. Chen, P., Dhanasobhon, S. and Smith, M.D. (2006) All Reviews are Not Created Equal, Working Paper.
7. Chevalier, J. and Mayzlin, D. (2006) The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research*, 4, 3, 345-354.
8. Dellarocas, C., Awad, N.F. and Zhang, X. (2004) Exploring the Value of Online Product Ratings in Revenue Forecasting: The Case of Motion Pictures, in *Proceedings of the 25th International Conference on Information Systems (ICIS)*, Washington, D.C., 379-386.
9. Duan, W., Gu, B. and Whinston, A.B. (2008) Do Online Reviews Matter? – An Empirical Investigation of Panel Data, *Decision Support Systems*, 45, 4, 1007-1016.
10. Forman, C., Ghose, A. and 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.
11. Ghose, A. and Ipeiritis, P.G. (2007) Designing Novel Review Ranking Systems: Predicting Usefulness and Impact of Reviews, in *Proceedings of the 9th International Conference on Electronic commerce*, Minneapolis, MN, 303-310.
12. Ghose, A. and Ipeiritis, P.G. (2008) Estimating the Socio-Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics, Working Paper.
13. Gupta, S. (1988) Impact of Sales Promotions on When, What, and How Much to Buy, *Journal of Marketing Research*, 25, 4, 342-355.
14. Hansen, M.T. and Haas, M.R. (2001) Competing for Attention in Knowledge Markets: Electronic Document Dissemination in a Management Consulting Company, *Administrative Science Quarterly*, 46, 1-28.
15. Hunt, R.E. and Newman, R.G. (1997) Medical Knowledge Overload: A Disturbing Trend for Physicians, *Health Care Management Review*, 22, 70-75.
16. Jeppesen, L.B. and Frederiksen, L. (2006) Why Do Users Contribute to Firm-Hosted User Communities? *Organization Science*, 17, 1, 45-64.
17. Lerner, J. and Tirole, J. (2002) Some Simple Economics of Open Source, *The Journal of Industrial Economics*, 50, 2, 197-234.

18. Li, X. and Hitt, L. (2008) Self Selection and Information Role of Online Product Reviews, *Information Systems Research*, Forthcoming.
19. Liu, Y. (2006) Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue, *Journal of Marketing*, 70, 3, 74-89.
20. Mayzlin, D. (2006) Promotional Chat on the Internet, *Marketing Science*, 25, 2, 155-201.
21. Ocasio W. (1997) Towards an attention-based view of the firm, *Strategic Management Journal*, 18, summer special issue, 187-206.
22. Pavlou, P. and Gefen, D. (2004) Building Effective Online Marketplaces with Institution-based Trust, *Information Systems Research*, 15, 1, 37-59.
23. Postmes, T., Spears R., Lea M. (2000) The formation of group norms in computer-mediated communication, *Human Communication Research*, 26, 3, 341-371.
24. Resnick, P., Zeckhauser, R., Friedman, E. and Kuwabara, K. (2000) Reputation Systems, *Communications of The ACM*, 43, 12, 45-48.
25. Reuters (1998) *Out of the Abyss: Surviving the Information Age*, London: Reuters Limited.
26. Scoble, R., and Israel, S. (2006) *Naked conversations: How Blogs are Changing the Way Businesses Talk with Customers*, John Wiley.
27. Simon, H.A. (1997) Designing Organizations for an Information-Rich World, In D.M. Lambertson (ed.), *The Economics of Communication and Information*, Cheltenham, UK, 187-203.
28. Zhang, X. and Dellarocas, C. (2006) The Lord of the Ratings: How a Movie's Fate is Influenced by Reviews?, in *Proceedings of the 27th International Conference on Information Systems (ICIS)*, Milwaukee, WI, 1959-1978.
29. Zhang, X., Dellarocas, C. and Awad N.F. (2004) Estimating Word-of-Mouth for Movies: The Impact of Online Movie Reviews on Box Office Performance, Paper presented at the *2004 Workshop on Information Systems and Economics (WISE)*, College Park, MD.

APPENDIX A

To study whether the reviewers are repeating previously mentioned product features and attitudes, we apply text mining technique to the review content. Following Ghose and Ipeiroitis (2007) and Ghose and Ipeiroitis (2008), we define nouns to be product features that are used in the review content and adjectives to be reviewers' attitude. The measurement is the frequency of commonly used words for individual review content.

When calculating the frequency, we only consider the reviews for the same book rather than use all data together. This is because the product features of a book are very different from those of digital products, such as digital cameras, which share very similar product features across different brands and models. The product features of a book are mainly the story, the characters, or the writing skills which vary a lot across books. Therefore, it makes more sense to study review content within each book than across all books. Otherwise, we may not be able to find interesting common words that are frequently used.

To obtain the frequency of commonly used words, we use the following three-step approach. First, we use SAS Text Miner to generate a term list for all the reviews of one book. This is a list of nouns and adjectives that have been used in the reviews. Then, we count the frequency of each term in the term list for each review. Finally, we calculate the variable, *Frequency*, which is the dependent variable to measure the frequency of commonly used words. Although there are usually over 300 hundreds of words in the term list, we are only interested in words that are frequently used. Therefore, we only consider the words that appear in at least 10 reviews and combine the count of frequency for these words for each review. Since the frequency will be affected by the length of each review, we divide the sum of count by the length of the review, which is the total number of words used in each review. This allows us to compare between reviews. In addition, since the number of such frequently used words may be different across different books, we further divide the frequency by the number of frequently used words. This allows us to compare reviews across books.

$$Frequency = \frac{\sum \text{Count of frequently used words}}{\text{Review Length} * \text{Number of frequently used words}}$$