Online Product Opinions: Incidence, Evaluation and Evolution

Wendy W. Moe  
Associate Professor of Marketing  
Robert H. Smith School of Business  
University of Maryland  
3469 Van Munching Hall  
College Park, MD 20472  
(301) 405-9187  
wmoe@rhsmith.umd.edu

David A. Schweidel  
Assistant Professor of Marketing  
Wisconsin School of Business  
University of Wisconsin – Madison  
975 University Avenue  
Madison, WI 53706  
(608) 265-2736  
dschweidel@bus.wisc.edu

December 2009

The authors would like to thank the Wharton Interactive Media Initiative and the Marketing Science Institute for supporting this research. The authors would also like to thank BazaarVoice for providing the data used in this research.
Online Product Opinions: Incidence, Evaluation and Evolution

ABSTRACT

While recent research has demonstrated the impact of online product ratings and reviews on product sales, we still have a limited understanding of the individual’s decision to contribute these opinions. In this research, we empirically model the individual’s decision to provide a product rating and investigate factors that influence this decision. Specifically, we consider how previously posted opinions in a ratings environment may affect a subsequent individual’s posting behavior, both in terms of whether to contribute (incidence) and what to contribute (evaluation), and identify selection effects that influence the incidence decision and adjustment effects that influence the evaluation decision. Systematic patterns in these behaviors have important implications for the evolution of product opinions at a site.

Our results indicate that individuals vary in their underlying behavior and their reactions to the product ratings previously posted. We demonstrate the implications of these behaviors for the evolution of product opinions through the use of simulations. We show that posted product opinions can be affected substantially by the composition of the underlying customer base and find that products with polarized customer bases may receive product ratings that evolve in a similar fashion to those with primarily negative customers as a result of the dynamics exhibited by a core group of active customers.
INTRODUCTION

The post-Internet marketplace is no longer limited to the one-way communications from sellers to buyers. Instead, consumers have become much more active in influencing and altering the nature of conversations around brands and products. Facilitated by developments in online technologies, consumers can easily contribute their thoughts and opinions to the marketplace through discussion groups, product ratings and reviews, and blogs. As a result, consumers have begun to talk with each other on a scale larger than marketers have previously experienced. However, this new environment is not without risks for marketers. In particular, marketers are increasingly losing control over the dialogue taking place around their products and brands. While this can be a positive development as consumers become more engaged and generate an increased level of buzz in the market, it can also have adverse consequences if the tone and content turn negative.

Of even more concern is that extant research has shown the existence of systematic biases in consumer product ratings. Several researchers have shown empirically that posted product ratings and reviews become increasingly negative as ratings environments mature (Li and Hitt 2008; Godes and Silva 2009). Schlosser (2005) also showed in a lab environment that posters adjust their product evaluations depending on the opinions expressed by others. As these studies demonstrate, a ratings environment can take on a life of its own, sometimes to the detriment of the product or brand to which it is dedicated. In some cases, the posted content provides a fair evaluation of the product/brand. However, as illustrated in the aforementioned studies, posted content can also reflect the influence of others.

The primary objective of this research is to empirically examine the behavior of individuals providing product ratings in an effort to better understand how expressed opinions systematically evolve over time. As part of this effort, we investigate the role that others’ ratings can have in influencing posting behavior. We consider two separate effects that may influence the subsequent evolution of product opinion. First, previously posting ratings may affect the incidence with which individuals choose to contribute their own opinions, which we refer to as a selection effect. Second, while some customers may prefer to provide their comments such that they will stand out from the crowd, others may prefer to
be consistent with the majority. Thus, in addition to selection effects, there may also be adjustment effects where individuals who ultimately decide to post may opt to revise their evaluations upward or downward based on previously posted comments. In other words, the ratings environment (as characterized by previously expressed opinions) may affect the composition of the posting population by systematically encouraging or discouraging certain types of posters as well as influencing the content that each contributor chooses to post. To disentangle these two forces, we therefore jointly model both the decision of whether to contribute (posting incidence) and what to contribute (posted evaluation) as a function of the ratings environment.

In contrast to much of the existing research in the area of online ratings and reviews, which has examined posted content at the product level, we model posted ratings at the level of the individual consumer. Many individuals actively contribute their opinions for a variety of different products. Examining an individual’s behavior across products, and hence ratings environments, can provide significant insights into the behavioral drivers of posting behavior. By jointly modeling posting incidence and the posted evaluation, we examine the impact that previously posted ratings may have on each aspect of posting behavior. In doing so, we separate the effect on the composition of the posting population from that on an individual poster’s stated evaluation, and provide a nuanced explanation for the negative trend in product ratings that has been observed in other studies.

Our findings show that there is substantial heterogeneity in individuals’ underlying incidence and evaluation behaviors. We also find heterogeneity in how individuals respond to the previously expressed opinions in a ratings environment. Our results reveal that highly active posters tend to be more negative in their evaluations and are more prone to post in dissentious environments. Furthermore, when they post to these environments, they post ratings that are differentiated from existing ratings. In contrast, less active posters tend to be more positive and are likelier to contribute to environments that exhibit a consensus of opinions. When these individuals post, they tend to exhibit “bandwagon” behavior by posting opinions that are consistent with the existing consensus.
In this research, we profile different types of posters and characterize them according to their incidence and evaluation behaviors. We then simulate a number of ratings environments that vary in terms of the composition of the product’s underlying customer base to illustrate trends in product opinion over time. In general, less active posters with positive opinions tend to contribute early in the process. As the ratings environment evolves, increasingly negative opinions from more active individuals become more prevalent.

Through simulations, we explore the effects of customer base composition on the evolution of product opinion and find that the behavior of a core group of active customers can significantly shape the direction of expressed opinions. Specifically, our results show that a highly polarized customer base will result in more negative opinions and exhibit a stronger downward trend than a more neutral customer base with the same median opinion. The ratings resulting from a polarized customer base tend to be dominated by customers with extremely negative opinions that are not representative of the entire customer base. Interestingly, the resulting ratings environment is not differentiable from that resulting from an exclusively negative customer base, highlighting the caution that must be exercised in drawing conclusions from a cursory view of a trend in expressed opinions.

In the next section, we discuss various factors that influence an individual’s posting behavior, both in terms of their posting incidence decision and their posted evaluation. This has direct implications for the composition of the posting population and hence posted ratings. We then present a model that captures the incidence and evaluation decisions simultaneously. Using the empirical results of this model, we then conduct a series of simulations to better understand how product opinions evolve and the effect of customer base composition on this evolution.

**WHY DO PEOPLE POST?**

The decision to engage in word-of-mouth activities has been extensively studied in offline environments. What makes online word-of-mouth unique is our ability as researchers to observe the behavior through online ratings and reviews. Dichter (1966) first proposed a set of motivational
categories that described why consumers engaged in the transmission of word-of-mouth. These categories have been refined and further explored by later researchers (Westbrook 1987). While many variations of these categories exist, a number of motivational themes are persistent. Some common motivations for transmitting word-of-mouth include self-enhancement (where the individual is interested in drawing attention to him/herself), altruism (where the individual is interested in helping others) and product involvement (where the individual derives an inherent utility in talking about the product). As the online environment evolved, these constructs have been extended to online word-of-mouth (Hennig-Thurau et al. 2004).

It seems obvious that there are differences across individuals in terms of posting motivation. However, these differences also have implications for how individuals’ posting behaviors are likely to respond to environmental factors, such as the product opinions previously posted.

What Influences Participation?

An individual’s posting behavior can be decomposed into two decisions: incidence (the decision of whether to post) and evaluation (the decision of what to post). Both of these decisions may be influenced by the individual’s independent evaluation of the product. Not only may an individual’s independent evaluation of a product affect his/her decision to contribute a rating for the product, but it would also logically affect the rating that is ultimately contributed. In the offline environment, Anderson (1998) showed that individuals with extremely negative opinions are more likely to engage in word-of-mouth activities. However, in the online environment, several studies have shown that posted product ratings have been overwhelmingly positive (see Chevalier and Mayzlin 2006 for an example). As a result, Dellarocas and Narayan (2006) suggest that individuals with extremely negative and individuals with extremely positive opinions will be more likely to post online than those with more moderate opinions.

In addition to the individual’s independent evaluation of the product, posting behavior is also subject to opinion dynamics. At the aggregate level, several studies have shown a downward trend in
posted product ratings as the ratings environment matures (Godes and Silva 2009; Li and Hitt 2008). In a controlled experimental setting, Schlosser (2005) shows that an individual poster has a tendency to adjust his/her posted product evaluation after viewing what others have posted. These studies together suggest that the ratings environment can lead to substantial ratings dynamics. Specifically, it has been argued that some posters try to differentiate themselves from others by posting more negative opinions (Schlosser 2005, Godes and Silva 2009), resulting in the downward trend in ratings as the environment becomes more populated.

While ratings dynamics is a relatively new area of research in the marketing literature, opinion dynamics is an established area of research in the political science literature. Specifically, political scientists have been studying the effects of opinion polls on election day behavior (see McAllister and Studlar 1991 for a review). Studies using both aggregate and individual-level data have shown that published opinion polls can affect voter turnout and election results. The direction of these effects have been the subject of extensive discussion, and researchers have debated the presence of a “bandwagon” effect, where opinion polls influence voter behavior in favor of the candidate leading in the polls (McAllister and Studlar 1991, Marsh 1984), versus an “underdog” effect, where the candidate trailing in the polls is favored (Gartner 1976, Straffin 1977). Still other researchers have shown that the declaration of a clear winner in opinion polls can depress overall voter turnout as voters perceive their votes to be inconsequential (e.g. Epstein and Strom 1981; Dubois 1983; Jackson 1983; Delli Carpini 1984; Sudman 1986). Overall, these studies show that knowledge of others’ opinions can affect both the individual’s incidence (voter turnout) and evaluation (voting choice) decisions.

In this paper, we will therefore examine an individual’s incidence and evaluation decisions as a function of both the individual’s underlying utility for the product as well as the ratings environment, as characterized by the posted opinions of others (see Figure 1). The covariates used in our model will allow us to identify a variety of potential effects suggested by the literature. We refer to factors that influence the incidence decision as having a selection effect and those that influence the evaluation decision as
having an *adjustment* effect. We examine possible differentiation effects, bandwagon effects and the effects of consensus (or dissention).

**Figure 1. What influences the posting decision?**

---

The Composition of the Posting Population

Figure 2 illustrates a process by which individuals are funneled through a series of decision stages before a comment is posted. When the ratings environment affects the evaluation decision, the poster may adjust his/her product rating in an effort to differentiate from the existing ratings (differentiation effect) or to conform to the existing ratings (bandwagon effect). This *adjustment effect* has been hypothesized to be the source of ratings dynamics by previous researchers (Schlosser 2005; Godes and Silva 2009). However, the ratings environment can also affect the incidence decision. That is, an individual may choose not to post at all as a result of the ratings environment. As a result, this *selection effect* can systematically influence the composition of the posting population. While previous empirical research on ratings dynamics has documented the adjustment effect, it has not simultaneously considered the selection effect.
As the framework presented in Figure 2 illustrates, posted opinions depend not only on the ratings dynamics, but also on the composition of the customer base. Though the full buying population may not engage in posting behavior, those customers who choose to contribute are drawn from this larger population. Consequently, characteristics of a product’s customer base can determine the composition of the posting population and may subsequently shape the evolution of product opinions in the ratings environment.

In this paper, we examine individual posting behavior and potential selection and adjustment effects. We present a model that captures the incidence and evaluation decisions at the individual-level allowing us to describe each individual in our data set according to their baseline incidence and evaluation behavior as well as their sensitivity to selection and adjustment effects. We then use the results of this model to simulate a variety of customer bases, each composed of a different subset of individuals that
vary according to their incidence and evaluation behaviors, in an effort to investigate the impact of
customer base composition on how expressed opinions evolve over time.

DATA

The data, provided by BazaarVoice, is from an online retailer of bath, fragrance and home
products and contains posted product ratings for six months in 2007 when the ratings functionality was
first introduced on the site. In this rating period, 4974 unique individuals posted product ratings, resulting
in a total of 10,460 ratings across 1811 products. Approximately 18% of the raters posted ratings for
multiple products. The data also indicate the time at which ratings were posted, facilitating identification
of the set of products for which an individual provides evaluations in each of his/her rating sessions.

Our objective is to model the behavior of each individual across a variety of products. With this
data set, this would necessitate the construction of a 4974 x 1811 matrix of ratings. Because of the
computational constraints associated with a matrix of such size, we sample 200 products from this data
set. However, to maintain a sufficient number of observations for each individual rater in the sample, we
draw a systematic sample as follows. We include the 100 most rated products at the site, which provide a
large base of individual raters, many of whom post ratings for multiple products. To ensure variation
across products and hence ratings environments, our sample also includes 100 additional products that
were chosen at random. Our sampling results in a dataset that includes 2436 individual raters who
provide a total of 3681 product ratings.

Covariates

Extant research in the ratings literature has converged on a set of metrics that best describe
previously posted ratings (see Dellarocas and Narayan 2006). These metrics have focused on the valence,
variance and volume of posted product ratings. Valence is typically represented by average rating;
variance has been measured using statistical variance measures as well as other dispersion measures such
as entropy; and volume is simply captured as the number of posted product ratings. However, the
interpretation of these metrics can sometimes be difficult, especially when each metric is treated as independent. Consider a ratings environment with a single 5-star rating compared to another with multiple 5-star ratings. The valence and variance of reviews in both cases are the same. The only measure differentiating the two ratings environments is the volume of ratings, and any differences in behavior between the two would be attributed to a volume effect. However, this volume effect likely interacts with the valence of the ratings. That is, a ratings environment with multiple 5-star ratings may be perceived as more positive than one with just a single 5-star rating while a ratings environment with multiple 1-star ratings may be perceived as more negative than one with just a single 1-star rating.

For our analysis, we consider the valence, variance and volume of previously posted ratings as well as the interaction effects among these three metrics. For parsimony and to eliminate for collinearity in these six descriptors, we performed a factor analysis on the set of 36,600 daily ratings environments (200 products x 183 days) as described by the main effects and interactions. The factor analysis results in two underlying constructs that explain 92% (61% by the first factor and 31% by the second factor) of the observed variation among the daily ratings environments (see Table 1). The first factor (F1) is strongly related to the variance of posted ratings while the second factor (F2) is influenced by both the valence and volume of ratings. In other words, F1 reflects the amount of consensus or dissention in the ratings environment while F2 captures the overall positivity of posted product ratings, taking into account the number of ratings that have contributed to this positivity. For the remainder of this paper, we use these factors as our model covariates rather than the raw ratings metrics.

<table>
<thead>
<tr>
<th>Table 1. Rotated component matrix resulting from factor analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
</tr>
<tr>
<td>------------------------------------</td>
</tr>
<tr>
<td>Valence</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Volume</td>
</tr>
<tr>
<td>Valence x Variance</td>
</tr>
<tr>
<td>Valence x Volume</td>
</tr>
<tr>
<td>Variance x Volume</td>
</tr>
</tbody>
</table>
MODEL DEVELOPMENT

Our modeling framework is similar to that of Ying, Feinberg and Wedel (2006). In their research, the authors predict individual consumer ratings for a selection of movies by modeling the rating incidence as a binary probit process and the rating evaluation as an ordered probit process, allowing for a correlation between the two processes. To examine the effects that ratings environments may have on how posted ratings evolve, we model an individual’s set of incidence and evaluation decisions in each of his/her ratings sessions, extending this modeling framework to (1) include covariates to capture the effects of the ratings environment and (2) condition the incidence model on the fact a rating is provided for at least one product in each rating session. Furthermore, we model the incidence and evaluation models as related processes by incorporating an individual’s utility for a product into both model components and by correlating the impact of covariates across the two model components.

Incidence Model: Whether or Not to Rate a Product

We model an individual’s decision to submit a rating for a product based on four components. First, individuals may vary in their tendency to submit product ratings. Whereas some may frequently submit ratings for products, others may be more judicious in deciding for which products to submit a rating. Second, the decision may depend on the product itself, as individuals may not consider posting evaluations for all products with the same tendency. The third consideration in the incidence model is an individual’s utility for a particular product. While individuals may be more prone to post ratings for those products for which they have a high utility, they may also be likely to post ratings for those products for which they have a low utility. Lastly, central to this research, the incidence decision may depend on the current state of the ratings environment at the time of the ratings session.

To incorporate these components into the incidence decision in a given session, we begin by assuming that individual i’s utility for product j is given by:

\[
U_{ij} = \gamma_{ii} + \kappa_j
\]
where $\gamma_i$ captures $i$’s baseline utility for products and $\kappa_j$ reflects variation in utility across different products, such that $\kappa_j \sim N(0, \sigma^2)$. Our specification for $U_j$ thus allows for variation across both respondents and across products in the utility function. We define the incidence utility for product $j$ on $i$’s $k^{th}$ rating session as $IU_{ijk}$, where:

$$IU_{ijk} = \beta_{1i} + \beta_{2i}F_{jk1} + \beta_{3i}F_{jk2} + \delta_1U_{ij} + \delta_2U_{ij}^2.$$  

The first term, $\beta_{1i}$, allows for variation in the propensity for an individual to submit a product rating. The terms $\beta_{2i}$ and $\beta_{3i}$ capture the effect of the current ratings environment on $i$’s decision of whether or not to submit a product rating. The coefficients $\delta_1$ and $\delta_2$ allow for variation in the incidence utility based on the $i$’s utility for product $j$. As we will see, these terms allow us to link the incidence model and the evaluation model in a flexible fashion.

To allow for variation in individuals’ consideration of posting ratings for different products, we assume that the probability that an individual considers submitting a rating for product $j$ is given by:

$$P(\text{consider}_{ij} = 1) = \Phi(\theta + \tau_j).$$

where $\Phi(\cdot)$ denotes the standard normal c.d.f., $\theta$ reflects the average tendency (across products) for customers to consider posting an evaluation, and $\tau_j \sim N(0, \sigma^2)$ to account for variation across products.

Conditional on considering submitting a rating for product $j$, the probability with which $i$ submits a rating for $j$ on his/her $k^{th}$ rating session is given by:

$$P(z_{ijk} = 1 \mid \text{consider}_{ij} = 1) = \Phi(IU_{ijk}).$$

where $z_{ijk}=1$ if $i$ posts a rating for product $j$ on his/her $k^{th}$ rating session and $z_{ijk}=0$ otherwise. The probability with which $i$ submits a rating for product $j$ on his/her $k^{th}$ rating session is then given by the product of equations (3) and (4). Thus, for $i$ to submit a rating for product $j$ ($z_{ijk}=1$), $i$ must have considered posting a rating for the product and $i$ was not deterred from doing so by the current state of the ratings environment. In the event that $i$ chose not to post a rating for product $j$ on his/her $k^{th}$ session ($z_{ijk}=0$), it may have occurred either because $i$ did not consider submitting a rating or because $i$ considered doing so but was dissuaded by the ratings environment.
We develop our incidence model at the level of the rating session. In order to observe a rating session, at least one product rating must occur; that is, it is not possible for \( z_{ijk} = 0 \) for all \( j \). The likelihood of observing the vector of posting decisions \( z_{ik} \) is then given by:

\[
\Pr(z = z_{ik}) = \frac{\prod_{j} \Pr(z_{ijk} = 1) \prod_{j} \Pr(z_{ijk} = 0)}{1 - \prod_{j=1} \left(1 - \Pr(z_{ijk} = 0)\right)}.
\]

Evaluation Model: What to Rate a Product

Should \( i \) choose to rate product \( j \) on his/her \( k^{th} \) rating session, the state of the ratings environment may also affect his/her expressed evaluation of the product. For example, some individuals may see consistently high product evaluations and choose to revise their evaluation upward so as to be consistent with others (bandwagon effect). On the other hand, given a consistently positive ratings environment, some may revise their evaluations downward so as to stand out from the crowd (differentiation effect).

We model an individual’s rating using an ordered probit model. We define \( i \)'s evaluation utility for product \( j \) on his/her \( k^{th} \) rating session as:

\[
EU_{ijk} = U_{ij} + \gamma_{2i}F_{jk1} + \gamma_{3i}F_{jk2}.
\]

The evaluation utility thus depends on \( i \)'s utility for product \( j \) (reflected by \( U_{ij} \)), as well as the effects of the ratings environment via the coefficients \( \gamma_{2i} \) and \( \gamma_{3i} \). As the current state of a ratings environment may impact both the decision to post a rating and the value of the rating, in contrast to Ying, Feinberg and Wedel (2006), we consider the correlation that may exist among the individual-level response parameters \( \beta \) and \( \gamma \) by assuming that:

\[
\begin{bmatrix}
\beta_i \\
\gamma_i
\end{bmatrix} \sim MVN \left( \begin{bmatrix}
\beta \\
\gamma
\end{bmatrix}, \Sigma \right)
\]

Ratings are submitted on a 5-point scale. Based on the evaluation utility, the probability that \( i \) submits a rating of \( y_{ijk} \) is given by:
To parameterize the cutpoints for the ordered probit model, we assume that $\mu_{1i} = \exp(\eta_{1i})$ and that $\mu_{ni} = \mu_{(n-1)i} + \exp(\eta_{ni})$ for $n=2$ and 3. We assume that the vector $\eta_i$ is drawn from a multivariate normal distribution.

In the spirit of Ying, Feinberg and Wedel (2006), we have presented a model for individual’s product evaluations, taking into account an individual’s decision to not provide ratings for many products on each ratings session. Like Ying, Feinberg and Wedel (2006), we allow the incidence and evaluation components of the model to be linked by assuming the product utility $U_{ij}$ affects both components. Our parameterization of this “utility link” allows for a non-monotonic relationship between the incidence utility and evaluation utility. While our framework allows for a monotonic relationship, we anticipate a U-shaped relationship in which both high and low levels of product utility increase the incidence of product postings ($\delta_2 > 0$).

Estimation

To fit the proposed model, we use a hierarchical Bayes procedure. Diffuse normal priors were assumed for the mean effects of covariates in the incidence and evaluation models ($\beta$ and $\gamma$, respectively), the mean tendency to consider posting an evaluation ($\theta$), and the effect of utility on the incidence decision ($\delta_1$ and $\delta_2$). For $\Sigma$, we employ an Inverse-Wishart prior (with diffuse priors on the hyperparameters) and directly sample the parameters. To make inferences under the proposed model, an MCMC sampler was run for 10,000 iterations which served as a burn-in period. We then obtained inferences from posterior samples from the next 20,000 iterations.
RESULTS

Table 2 provides the parameter estimates from our model estimation. We present the mean and standard deviation (across individuals) for $\beta$, $\gamma$, and $\mu$, which indicate substantial heterogeneity across individuals in terms of how they respond to different aspects of the ratings environment. Focusing first on the incidence model, we observe noticeable variation across individuals in terms of their preference for posting in consensus. While some may be more prone to provide a product rating when previous posters were in agreement ($\beta_2<0$), others may abstain from posting in such environments and instead exhibit a preference for contributing product ratings when such a consensus has not yet been reached ($\beta_2>0$). In contrast to the variation that we observe across individuals in how they respond to consensus (or the lack thereof), virtually all individuals exhibited a preference for posting in positive ratings environments ($\beta_3>0$). We illustrate the distribution of posterior means across individuals from the incidence model in Figure 3.

We next turn our attention to the evaluation model. Figure 4 provides a histogram of posterior estimates related to the evaluation model. In contrast to the incidence model in which individuals responded differentially to the level of consensus in the ratings environment, we see that the variance of expressed opinions has a limited effect on the rating an individual provides. However, the results indicate that the overall positivity of the ratings environment has varying effects on what an individual posts. Though some adjust their rating upward as the positivity of posted ratings increases ($\gamma_3>0$), and hence provide higher evaluations when previous posters have already done so, we see that others adjust their rating downward as the positivity increases ($\gamma_3<0$), and consequently lower their reported product evaluations. We thus observe substantial heterogeneity in individuals’ propensities to exhibit bandwagon versus differentiation effects in the evaluation stage.
<table>
<thead>
<tr>
<th>Component</th>
<th>Parameter</th>
<th>Description</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_j$</td>
<td>Product-level probability of</td>
<td>.02 (.01)</td>
</tr>
<tr>
<td>Incidence Model</td>
<td>$\beta_{1i}$</td>
<td>considering posting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_{2i}$</td>
<td>Incidence intercept</td>
<td>-2.48 (.167)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{3i}$</td>
<td>Effect of $F_1$</td>
<td>.02 (.18)</td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>Linear “utility link”</td>
<td>.53 (.24)</td>
</tr>
<tr>
<td></td>
<td>$\delta_2$</td>
<td>Quadratic “utility link”</td>
<td>.06 (.01)</td>
</tr>
<tr>
<td>Evaluation Model</td>
<td>$\gamma_{1i}$</td>
<td>Mean product utility</td>
<td>3.05 (.49)</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{2i}$</td>
<td>Effect of $F_1$</td>
<td>-.05 (.09)</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{3i}$</td>
<td>Effect of $F_2$</td>
<td>-.06 (.24)</td>
</tr>
<tr>
<td></td>
<td>$\log(\mu_{1i})$</td>
<td>Difference in cutoff for</td>
<td>-.115 (.12)</td>
</tr>
<tr>
<td></td>
<td>$\log(\mu_{2i}-\mu_{1i})$</td>
<td>Rating=2</td>
<td>-.82 (.06)</td>
</tr>
<tr>
<td></td>
<td>$\log(\mu_{3i}-\mu_{2i})$</td>
<td>Rating=3</td>
<td>-.63 (.12)</td>
</tr>
<tr>
<td></td>
<td>$\log(\mu_{4i}-\mu_{3i})$</td>
<td>Rating=4</td>
<td></td>
</tr>
</tbody>
</table>
In addition to the state of the ratings environment at the time of a ratings session, an individual’s decisions may be influenced by his utility for the product, $U_{ij}$. While it is tautological that an individual’s utility for a product will affect his/her evaluation of it, $U_{ij}$ may also affect the incidence decision. Figure 5 shows the relationship between utility (U) and incidence utility (IU), as governed by the parameters $\delta_1$. 

**Figure 3.** Posterior means from incidence model

**Figure 4.** Posterior means for evaluation model
and $\delta_2$. We see that individuals are more likely to submit product ratings when their product utility is high, which is consistent with empirical findings in the online ratings literature that shows a strong positivity bias in online product ratings (Dellarocas and Narayan 2006). We also observe an increase in the likelihood of posting incidence when utility is low, consistent with the findings of Anderson (1998) showing that consumers are more likely to engage in word-of-mouth activities when they are unsatisfied. This empirical finding illustrates the potentially non-linear relationship between the incidence and evaluation decisions, contrasting with the monotonic relationship that has been assumed in previous research (e.g., Ying, Feinberg and Wedel 2006).

**Figure 5. Role of product utility in rating incidence**

![Figure 5](image)

To further examine the interdependencies between the incidence and evaluation decisions, we present the correlation matrix between the incidence model parameters ($\beta$) and the evaluation model parameters ($\gamma$) in Table 3. The correlation coefficients indicate a number of interesting relationships, particularly with respect to the incidence model intercept, which reflects the frequency with which an individual posts a rating. We observe that frequent raters are more likely to post in dissident environments ($r = .42$) and more positive environments ($r = .48$). In terms of their evaluation behavior,
these frequent raters also tend to be more negative in their ratings \( r = -.77 \) and tend to differentiate their posted ratings from others \( r = -.32 \).

**Table 3. Correlation Matrix**

<table>
<thead>
<tr>
<th></th>
<th>Incidence Model</th>
<th>Evaluation Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta_1 )</td>
<td>( \beta_2 )</td>
</tr>
<tr>
<td>Intercept (( \beta_1 ))</td>
<td>1</td>
<td>( .42^* )</td>
</tr>
<tr>
<td>Variance (( \beta_2 ))</td>
<td>( .48^* )</td>
<td>( .37^* )</td>
</tr>
<tr>
<td>Valence-Volume (( \beta_3 ))</td>
<td>( .48^* )</td>
<td>( .37^* )</td>
</tr>
<tr>
<td>Intercept (( \gamma_1 ))</td>
<td>( -.77^* )</td>
<td>( -.19^* )</td>
</tr>
<tr>
<td>Variance (( \gamma_2 ))</td>
<td>( -.13 )</td>
<td>( .09 )</td>
</tr>
<tr>
<td>Valence-Volume (( \gamma_3 ))</td>
<td>( -.32^* )</td>
<td>( -.46^* )</td>
</tr>
</tbody>
</table>

* indicates 0 is not contained in the 95% HPD interval

To aid in understanding the emerging patterns in how individuals respond to the ratings environment in their incidence and evaluation decisions, Figure 6 provides an overview of the effect of the environment on the incidence decision. To more efficiently present the results, we group the posterior estimates into deciles (from lowest to highest) with each row in the matrix representing the response to the variance factor (F₁) and each column representing the response to the valence-volume factor (F₂). The elements within the matrix indicate the number of individuals who possess that combination of coefficients. We observe two clear clusters of individuals. While the cluster in the upper-left represents individuals who prefer posting in environments with a consensus and have weak positivity preferences, the cluster in the lower-right represents those who prefer posting in dissentious environments and have a relatively strong preference for high positivity environments.
We next consider how individuals’ preferences as to which environments to post in relate to their evaluation decisions. Table 4 categorizes individuals according to their preference for posting in ratings environments exhibiting consensus ($\beta_2 < 0$) versus dissent ($\beta_2 > 0$), and their preference for posting in positive environments ($\beta_3$) based on a median split. Again, the cluster of individuals in the upper-left quadrant prefer posting to environments with consensus and have no real valence-volume preference. These individuals tend to be the least frequent raters ($\beta_1 = -3.661$) and exhibit bandwagon effects when they do post ($\gamma_3 = .157$). We refer to these individuals as “low-involvement” posters because of their low frequency in contributing.

In contrast, the cluster of individuals in the lower-right quadrant tend to be fairly active raters ($\beta_1 = -1.398$) and exhibit differentiation behavior when they post ($\gamma_3 = -.230$). We refer to these individuals as “community-builders” primarily because of how active they are in contributing their opinions. Their tendencies toward more negative and differentiated postings are consistent with previous research showing that individuals who want to be perceived as more “expert” often try to differentiate themselves with more negative opinions (Schlosser 2005; Amabile 1983), as evidenced by the lowest mean value of $\gamma_3$ among these four clusters.
Table 4. Relationship between incidence and evaluation behavior

<table>
<thead>
<tr>
<th></th>
<th>Weak preference for high ValVol ($\beta_3&lt;$median)</th>
<th>Strong preference for high ValVol ($\beta_3&gt;$median)</th>
</tr>
</thead>
</table>
| Prefer consensus product forums ($\beta_2<$0)  | $\beta_1 = -3.661$  
  $\gamma_2 = -.037$  
  $\gamma_3 = .157$  
  Low-involvement | $\beta_1 = -2.513$  
  $\gamma_2 = -.104$  
  $\gamma_3 = .092$  
  Bandwagoners |
| Prefer dissensious product forums ($\beta_2>$0) | $\beta_1 = -2.627$  
  $\gamma_2 = .010$  
  $\gamma_3 = -.170$  
  Differentiators | $\beta_1 = -1.398$  
  $\gamma_2 = -.065$  
  $\gamma_3 = -.230$  
  Community-builders |

The remaining quadrants are less populated and are not significantly differentiated in terms of frequency of posting behavior. Instead, what differentiates the individuals in these quadrants is their tendency for bandwagon ($\gamma_3>$0) versus differentiation ($\gamma_3<$0) behavior in their posting evaluation decision. Overall, bandwagon and differentiation behaviors in the evaluation stage are correlated with the variance preference in the incidence stage. That is, individuals who prefer to post in environments with a consensus of opinions tend to exhibit bandwagon behavior when they post, whereas individuals who prefer to post in dissentious environments tend to exhibit differentiation behavior when they post.1

The incidence and evaluation models focus on the quantitative product ratings that individuals provide and shed light on how individuals respond to ratings environments. To gain additional insights into the motivations of these individuals, we examine the textual content of their product reviews. As mentioned earlier, extant research has converged on three motivations for engaging in word-of-mouth activities: (1) self-enhancement, (2) altruism and (3) product-involvement. Therefore, we analyze posted review text for each of these three motivations. Reviews that use the words me, I, we and other words

1 Recognizing the pitfalls associated with dichotomizing continuous measures (Irwin and McClelland 2003), we considered a continuous representation of these effects. The results reveal the same patterns observed in Table 4 and are available from the authors upon request.
that reference the first person are identified as “self-enhancing.” Reviews that use the words you, your and other words that reference the second person are identified as “altruistic.” All others are considered to be a result of a high “product-involvement” motivation.

Table 5 describes the posted review content for each of cluster of individuals identified in Table 4. Overall, a large proportion of reviews tend to reference oneself, and very little variation in this measure is observed across the different clusters. More interestingly, we observe that “community-builders” are significantly more likely to provide reviews that project an altruistic tone or focus exclusively on the product. This behavior is consistent with the interpretation of this highly active cluster as “community-builders” who are engaged with the product and have an altruistic desire to contribute helpful advice to the larger community.

<table>
<thead>
<tr>
<th>Table 5. Textual content of reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation Intercept ($\gamma_1$)</td>
</tr>
<tr>
<td>Low-involvement</td>
</tr>
<tr>
<td>Bandwagoner</td>
</tr>
<tr>
<td>Differentiator</td>
</tr>
<tr>
<td>Community-builder</td>
</tr>
</tbody>
</table>

Overall, the results presented in this section provide a description of how individuals behave in each stage of the rating process and how they respond to the ratings environment. As the clusters described in Tables 4 and 5 respond differently to ratings environments, the composition of the posting population changes, and the nature of the opinions being provided evolves. In the next section, we further investigate the implications of these behaviors for the overall ratings environment by simulating a number of ratings environments. These simulations will allow us to better understand how product opinions evolve and what factors influence this evolution.
THE EVOLUTION OF PRODUCT OPINION

In this section, we simulate ratings environments based on the results obtained from our model in an effort to more closely examine the drivers of observed evolutionary patterns. Our simulation procedure is as follows:

1. Based on the model parameters, we simulate a population of 10,000 individuals
2. From the population of 10,000 individuals, we draw a sample of 1,000 individuals to represent the product’s customer base.
3. We categorize each individual in the customer base as a community-builder (CB), low-involvement poster (LI), bandwagoner (BW) or differentiator (DF).
4. We compute the incidence probability for each individual in the customer base according to equation (4) and simulate posting incidence.
5. We compute evaluation utility for each poster according to equation (6) and simulate the posted rating according to equation (8).
6. We repeat steps 4 and 5 until 50 ratings are posted.
7. We repeat steps 1-6 for 5,000 iterations and average the results across iterations.

Simulation 1: Overall Evolutionary Patterns

For the first simulation, we randomly draw our customer base from the total population. Figure 7 plots the average rating and the variance in ratings over time while Figure 8 plots the participation of the various poster-types. As more ratings arrive, we see that the average rating decreases and the variance across ratings increases. The decomposition of the poster population in Figure 8 provides some insight as to what is driving this trend.\(^2\) Overall, community-builders are over-represented in the posting population in part because they are more active posters. However, they are relatively less active initially while low-

---

\(^2\) For clarity we present only the trends associated with the low-involvement population and the community-builders. The trend associated with bandwagoners is very similar to that seen among the low-involvement individuals while the trend associated with differentiators mirror that of the community-builders. These results are available from the authors upon request.
involvement customers are more active. As more ratings get posted, the ratings environment is likely to exhibit more variation in opinions. This increase in rating variance along with the increase in volume tends to encourage more community-builders to post. Since community-builders tend to be more negative in their evaluations, ratings gradually become more negative. Additionally, since community-builders exhibit differentiation behavior when posting a rating, the variance in ratings slowly increases. Though this trend has been commonly observed in ratings environments by past researchers (Li and Hitt 2008; Godes and Silva 2009), our analysis and simulation indicate that the source of this trend is a shift in the composition of the posting population. That is, the selection effect is central to the way in which expressed opinions evolve.

**Figure 7.** Simulated ratings evolution resulting from a representative customer base
Simulation 2: Diversity of Customer Opinions

Our model results indicate that an individual’s posting behavior is correlated with his/her underlying utility for the product. Therefore, we next compare a ratings environment resulting from a highly polarized customer base to one that is relatively homogeneous and representative of the median opinion. To construct the highly polarized customer base, we sample the 500 individuals with the lowest evaluation model intercept ($\gamma_1$) and the 500 individuals with the highest evaluation model intercept. To construct the median customer base, we sample the 1,000 individuals in the middle of the distribution. Note that the average opinion is similar across these two customer bases while the variance is dramatically different.

Before discussing the results of this simulation, it is important to understand the composition of these two different customer bases (see Table 6). The overall composition of poster types is very similar. However, substantial differences emerge when you more closely examine the differences between those who hold extremely positive versus extremely negative opinions in the polarized customer base (see
Table 7). Those with more negative opinions tend to be community-builders while those with more positive opinions tend to be low-involvement individuals. Since the community-builders are the more active posters, this will have significant implications for the ratings environment. Specifically, posted product ratings will be more negative, and positive opinions will be infrequent since those are held primarily by the low-involvement individuals.

**Table 6. Comparison of customer base composition**

<table>
<thead>
<tr>
<th></th>
<th>Representative</th>
<th>Median</th>
<th>Polarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-involvement</td>
<td>29.3%</td>
<td>28.8%</td>
<td>31.1%</td>
</tr>
<tr>
<td>Community-builders</td>
<td>32.8%</td>
<td>32.2%</td>
<td>34.3%</td>
</tr>
<tr>
<td>Bandwagoners</td>
<td>17.3%</td>
<td>17.7%</td>
<td>15.7%</td>
</tr>
<tr>
<td>Differentiators</td>
<td>20.7%</td>
<td>21.2%</td>
<td>18.9%</td>
</tr>
</tbody>
</table>

**Table 7. Comparison of negative versus positive customers in a polarized customer base**

<table>
<thead>
<tr>
<th></th>
<th>Extremely Negative</th>
<th>Extremely Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-involvement</td>
<td>13.9%</td>
<td>48.4%</td>
</tr>
<tr>
<td>Community-builders</td>
<td>52.6%</td>
<td>16.0%</td>
</tr>
<tr>
<td>Bandwagoners</td>
<td>17.8%</td>
<td>13.6%</td>
</tr>
<tr>
<td>Differentiators</td>
<td>15.7%</td>
<td>22.0%</td>
</tr>
</tbody>
</table>

Figures 9 and 10 provide the simulation results for the median customer base and the polarized customer base, respectively. As expected, the posted ratings resulting from the homogeneous, median customer base are relatively uniform. In contrast, the posted ratings from the polarized customer base are more negative and exhibit a strong downward trend over time (likely due to the differentiation behavior of the more active community-builders).
**Figure 9.** Simulated ratings evolution resulting from a median customer base

![Graph showing simulated ratings evolution for a median customer base.]

**Figure 10.** Simulated ratings evolution resulting from a polarized customer base

![Graph showing simulated ratings evolution for a polarized customer base.]
The comparison between these two customer bases is insightful. While the two customer bases share the same average opinion of the product (based on individuals’ values of $\gamma_1$), we see that the posted ratings are quite different. When the customer base is polarized, initial ratings will tend to come from the more negative community-builders. This leads to a dynamic where other community-builders enter the discussion and further differentiate by providing more negative opinions while low-involvement individuals refrain from posting in this increasingly dissentious environment. Eventually, posted ratings are dominated by extreme negative opinions that are not representative of the entire customer base. For comparison purposes, Figure 11 presents the simulated ratings resulting from an exclusively negative customer base. Interestingly, ratings generated by the polarized customer base are not differentiable from that of the negative customer base.

Figure 11. Simulated ratings evolution resulting from an exclusively negative customer base

Overall, the simulations show that the evolution of product opinions in a ratings environment can be heavily influenced by a core group of customers. In a more balanced customer base, ratings drift downward as community-builders enter. In a polarized customer base, posted opinions are dominated by a core group of more critical community builders. Their negativity and differentiation behavior has a
strong effect on ratings environment and as a result, posted product ratings do not necessarily represent
the opinions of the entire customer base.

**DISCUSSION AND CONCLUSIONS**

While previous work has studied the effects of online product reviews on consumers’ purchase
decisions at the product level, research to date has not explored the individual-level decisions of whether
to post an opinion or what to post. In this paper, we present a joint modeling framework to examine the
effects of previously posted content on posting incidence and evaluation decisions. While the evaluation
decision results in an adjustment effect that alters the content of future postings, the incidence decision
results in a selection effect that can shape the composition of the posting population. This latter dynamic
has been largely ignored in the extant research on online product opinions.

Our empirical analysis reveals that previously posted ratings can affect the tone that future
postings will take through both selection effects and adjustment effects. While both the variance and
overall positivity in a ratings environment impact the incidence decision, only the positivity of opinions
appears to influence the extent to which individuals revise their product evaluations when they post an
online product rating. Our results further suggest that the incidence and evaluation decisions are related.
Based on these behaviors, we identify four types of posters: (1) low-involvement individuals, (2)
community-builders, (3) bandwagoners, and (4) differentiators, and demonstrate systematic differences in
incidence and evaluation behaviors across these groups. Overall, we find that online opinions are
dominated by the more active community-builders. Furthermore, participation by community-builders
tends to increase over time while participation by low-involvement individuals tends to decrease. This
shift in the composition of the posting population can substantially affect the overall tone of posted
opinions.

Additionally, the composition of the customer base can exert a substantial influence on the
manner in which posted online opinions evolve. Due to selection and adjustment effects, the content
posted may not necessarily reflect the customer base’s overall opinion of the product. Rather, a vocal subset of the customer base may dominate the ratings environment, consequently steering the subsequent evaluations that are posted and deterring some customers from contributing to the environments. Marketers must consequently exercise caution in drawing inferences based on what they observe in their product reviews, as the opinions they observe may not provide an accurate gauge of the overall customer base’s perceptions.

There are a number of directions that remain open for future research. While we develop a multivariate incidence model for the set of products for which an individual posts evaluations in a single ratings session, research may examine the drivers of the time between sessions. Research may also explore methods to identify those posters who are the most influential, both in terms of steering the way in which online opinion evolves and in terms of affecting product performance. Though our empirical application considers product reviews, it may also be worthwhile to examine the way in which other types of product forums are shaped by such opinion dynamics. The modeling framework we present can be generalized for such analyses, taking into account the selection and adjustment effects.

Our research context has centered around the posting of product ratings, and as such the findings are limited to an individual’s one-time contribution to a given ratings environment. However, repeated contributions regarding a single product are common in discussion forums. While this repeat behavior is also important to understand, it lies outside the scope of this research. Future study of repeat posting behavior in the context of more flexible discussion forums can significantly add to our understanding of user-generated content and how this increasingly important factor in the consumer decision process is created.

As the online marketplace evolves and becomes increasingly interactive, consumers play a larger role in the creation of content that can influence the success or failure of a product. Because of this, it is critically important for marketers to understand why people contribute their opinions to online forums and what influences their behavior. This paper contributes to this larger effort.
REFERENCES


