Relative Importance of Product Rating and Amount of Rating to Consumer Choice

Lily Assaad, Nick Rector, Carey Cox

Abilene Christian University

Abstract

Extensive research has been conducted on the e-commerce concept of online reviews and their causes, but nothing has been done to discover the relative effect of ratings and amount of reviews on online consumer-choice, which we wanted to understand. We hypothesized that the customers' initial expectations of the product (based on past experience), product ratings, and product review-amounts would all combine to significantly predict the consumer's choice. However, the two independent variables, "rating" and "amount of reviews," alone, would not cause significant predictions.

Data collection took place at a campus-center booth where 48 interested participants received a sheet requiring choice between two identical chocolates, differing only in the amount of reviews and the averaged ratings, and which was made to look like the ratings system used by Amazon.com. Participants indicated their choice on the sheet and then received the chocolate. After data collection, a Bayesian mathematical model relating consumer-initial expectations with the two independent variables was created in order to predict forced consumer-choice from a binary choice-list.

We ran a multiple linear regression using rating, amount of reviews, and the Bayesian model as three independent variables. The ANOVA results showed that this model significantly predicted consumer choice. Additionally, the Bayesian model was the only significant predictor.

Results imply that companies should advertise high ratings more regularly. Additionally, amounts of reviews should be increased if ratings are high, and decreased if they are low. Current businesses do not practice this.

Introduction

Carey Cox is a college student at Abilene Christian University who often shops online, using websites such as amazon.com, barnesandnoble.com and macys.com, all three of which use a rating system for each item. While on amazon.com, she is considering the purchase of a new iPhone case. According to their descriptions, the two cases Carey is considering are similar in style, utility and price. However, one case has been given a four-star rating by seven different users, while the other case has a 3-star rating given by over two hundred users. Which of these cases will she choose?

With the rise of Internet came the rise of “e-commerce.” Retail, in an online environment, is potentially among the top usages of progressive information technology. Online interactions include an exchange of resources such as products, services and information, as well as more traditional advertising and customer support activities, which are now becoming extensively researched (Bruce, Moore & Birtwistle, 2004). One of the difficulties involved with electronic commerce is the ability to establish a sustainable competitive advantage. Because consumers can so easily compare prices and features across industries and individual products, achieving differentiation can be problematic, which places importance on customer reviews (Adam, 2011).

This online retail phenomenon ushered in the concept of “ratings and reviews,” which now plays a vital role in helping consumers make decisions. The implications of allowing unbiased user reviews to appear near a company’s product provide grounds for discussion of ethical and corporate value practices. If a company, such as Amazon, places value in the mutual respect and trust of customers, it must carefully consider allowing the customer experience to be announced in a public forum (Heydari, Jamehshooran and

Teimouri, 2011). The research community quickly reacted by exploring this topic. For example, it is now known that there is a temporal sequence within reviews (Godes and Silva, 2012). Moreover, online reviews are known to be more helpful when they are detailed, but that depends on the product type (Mudambi and Schuff, 2010). These user reviews may also be differentiated in strength of opinion and recommendation, as well as product aspect focus, based on users' culture (Chou, He, Lai and Zhou, 2013). Research has even been done regarding the *causes* of reviews, which turned out primarily to be other reviewers' comments, and social influence (Sridhar and Srinivasan, 2012).

However, on Amazon.com, one of today's largest online retailers, before a consumer can view all the product details, they are shown a screen with only five pieces of information (per product), for all the products searched: the product picture, product name, rating (from one to five), amount of ratings, and price. This can be seen in the image below of amazon's website (see Figure 1).

Thus, information on the rating and amount of ratings present comprises half of the total information any customer first views for each product they are exposed to. Yet, no research has been done to discover what the effect of these two numerical symbols has had on customer choice. Thus, we want to understand whether the *number* of ratings (independent variable I), as compared to the product rating (independent variable II), affects a consumer's choice (dependent variable), and if so, whether it has a larger or smaller effect on the decision than the product rating. We hypothesize that the number of ratings will have an effect, but that the rating itself will have a larger effect.

Methodology and Analysis

Data Collection

Since we planned on exploring the relative importance of the product rating versus the number of ratings given to the consumer's choice, the combination of the rating and amount of reviews for each product were the two independent variables, with the consumer's choice being the dependent variable. The term "product rating" refers to the rating out of five, which is created through averaging out all the individual consumer ratings for that product. Also, the "number of ratings" refers to the cumulative amount of online product ratings that people have submitted to the e-retailer.

This experiment was run in the campus center of a small, religiously affiliated, private university, and included 48 participants. We recruited them through setting up a booth in the campus center and allowing people to come explore how they could receive free chocolate. Thus, we used a convenience sampling method. Since the only thing the participants had to fill out was the experimental paper (which will be expounded on later), no other data regarding their age, gender or ethnicity was collected. The materials used for this experiment include paper, pencil and chocolates.

There were a possible five ratings and four amounts of reviews (i.e.: two variables, one with five levels and one with four levels). Yet, there were only 48 possible comparable binary choices. The experimenters manually paired the comparable choices, and had each individual pair of options printed on a sheet of paper (allowing the subject to choose out of two possible options). Thus, no two sheets of paper contained the same pair of choices. The format of the paper resembled an Amazon.com product layout, where the picture of the product was placed on the left side, with the right side containing the amount of stars and the number of reviews present. There was also a check box by each product. The product rating was a number out of 5 (shown in the form of stars), and the amount of reviews present ranged from 1-1000.

The experimenter explained that he would like the subject to choose which chocolate he/she would like him to give to him/her and handed the subject a piece of paper and pencil. The subject then indicated on the paper which of the two chocolates he/she wanted, and the experimenter gave him/her the chocolate of choice. We conducted this experiment in a very transient part of campus, where people do not stay for more than a few minutes. This was so that there would be a decreased chance of them telling each other about the experiment if they would have discovered that it was the same chocolate for either box that they could have ticked. Also, The experimenter running the experiment remained constant, and repeated the same phrase to all subjects, “would you like to have a free piece of chocolate? Indicate which one you would like me to give you on this sheet of paper.” This limited both experimenter bias and subject bias. The experimenter collected all 48 subject data-pieces within the same hour, limiting any results arising from a change of day or time within the data-collection period. The phrasing on all the sheets of paper was identical, with the only thing varying being the combination of rating with amount of reviews present for the product.

Behavioral Model

After data collection, a Bayesian mathematical model relating the two independent variables, online product rating and amount of reviews, was created in order to predict forced consumer-choice from a binary choice-list. This model also assumed that consumers in part make their choices based on their initial expectations of the products; these expectations are created through the consumers' past experience with similar products. Therefore, our model was based on the assumption that consumers combine their initial expectations of the product with the given online rating information in order to determine which option consumers would choose.

The inclusion of initial expectations within this Bayesian model was necessary because combining the customers' initial expectations with the given online ratings to produce the final assessment would actually give a very different result than if initial customer expectations were to be omitted. For example, if the number of online reviews is very small, a customer's initial expectations -- which can be thought of as internally-produced product reviews created based on past personal experience -- would influence their assessment of the product more than the online reviews. If, however, there were a very large number of online reviews on that particular choice-option, the customer's initial expectations would play a relatively miniscule role in affecting their final assessment of that product.

A mathematical instantiation of the above would be:

$$V = \frac{(i \cdot w) + (s_1 \cdot n_1)}{w + n_1} - \frac{(i \cdot w) + (s_2 \cdot n_2)}{w + n_2}$$

In the above formula, *V* represents the resultant numerical value indicating the customer-choice prediction. A positive value predicts customer choice of Option_1, whereas a negative value predicts choice of Option_2. The parameters affecting Option_i's final value are *b, w,* s_i, *and* n_i. The parameters pertaining to the customer's initial expectations are *i* and *w*, where *i* indicates the customer's initial star-rating of the option (the internal option-rating based on personal past experiences) and *w* indicates the weight that the initial rating has on the customer's final value of that option. In comparison to the number of online reviews submitted for a consumer product, *w* can be understood as the amount of reviews that are present *within* the customer's mind, based on past experiences that created the internal star rating. The parameters pertaining to the online rating and review are s_i, and n_i, where s_i indicates the online star-rating of Option_i, and n_i indicates the number of reviews submitted and used to assign that star-rating to Option_i.

Consider this example: A customer is being asked to choose between two free chocolate pieces, and is provided with the chocolates' rating and amount of reviews on Amazon.com. Sample_1 has a one-star rating and one review, where as Sample_2 has a two-star rating and 1000 reviews. Which option do you think they will choose? We will assume that the customer's initial expectation of these chocolate samples (based on his/her past experience with chocolate) is a three-star rating with a weight of approximately 50 reviews. We assume three stars by averaging the minimum and maximum amount of stars an option can receive ($\frac{1+5}{2}$). Inserting these values into the original equation would result in the following:

$$V = \frac{(i \cdot w)+(s_1 \cdot n_1)}{w+n_1} - \frac{(i \cdot w)+(s_2 \cdot n_2)}{w+n_2} = \frac{(3 \cdot 50)+(1 \cdot 1)}{50+1} - \frac{(3 \cdot 50)+(2 \cdot 1000)}{50+1000} = \frac{151}{51} - \frac{2150}{1050} = 0.913$$

Since the value of V turns out to be positive, the Bayesian model predicts that the customer would choose Sample_1 over Sample_2. Note, however, that a model using either rating or number of reviews alone would actually predict customer choice of Sample_2 instead.

In order to increase the formula's prediction accuracy, we needed to optimize parameter *w*. Thus, we hand tuned the *w* parameter by running repeated linear regressions on the formula until we found the value of w that maximized r^2, which turned out to be 40. Thus, the final Bayesian model is: $V = \frac{(i \cdot 40)+(s_1 \cdot n_1)}{40+n_1} - \frac{(i \cdot 40)+(s_2 \cdot n_2)}{40+n_2}$.

Results

We wanted to discover whether the general model could significantly predict choice or not, and we found out that it does. We discovered this through using a multiple linear regression model where we included three independent variables: difference in the two options' number of reviews (D_N), difference in the two options' star ratings (D_R), and the final Bayesian model value (V). The dependent variable was actual consumer choice.

This resultant model significantly predicted choice:

$$Y = 0.222 - 0.105D_R - 0.319V$$

In the above multiple linear regression, 0.222 is the constant; it represents the level of bias inherent in the equation as a whole. The variable D_R is obtained through subtracting the rating for $Option_2$ from the rating for $Option_1$. The variable V represents the resultant value of the Bayesian model explained in the Methodology section: $V = \frac{(i \cdot 40)+(s_1 \cdot n_1)}{40+n_1} - \frac{(i \cdot 40)+(s_2 \cdot n_2)}{40+n_2}$. The variable D_N did not prove to have any effect on the general model's ability to predict consumer choice, and was therefore not included by the data analysis in the general model.

The ANOVA for this general model yielded significant results (*$F(3,44)=6.872$, $p<0.001$*), showing that it could in fact predict consumer choice using the three variables. $R^2=0.319$, showing that the general model accounted for approximately 32% of variance within the data. Possible sources of variance could have resulted from the numerous uncontrolled variables present during data-collection.

We also wanted to know, all else being held equal, which, if any, of the independent variables significantly contributed to the regression model's predictability of choice. The only independent variable that significantly contributed was the Bayesian model value ($\beta = -0.319, t(47) = -2.88, p < 0.006$). In fact, the Bayesian model was able to accurately predict the final choice of 79.2% of the 48 participants. The variable D_R, when interacting with the other independent variables, had a large enough impact on the ability of the general model to predict choice to be included in it. However, statistically, it did not significantly contribute to the general model's power to predict ($\beta = -0.105, t(47) = -1.53, p < 0.134$).

Conclusion and Recommendations

The data shows that consumers' purchasing behavior is influenced by having comments and ratings of products present in the purchasing process. This find could be beneficial or detrimental for businesses, depending on which side of the scale they lie. By having a quality amount of good comments of their products, a business has much to gain by making the information easily accessible or even marketing the information. The opposite is true for businesses who receive numerous amounts of bad reviews for their products. Knowing that the relationship between the number of comments and average reviews is significant will encourage more competition and improving the purchasing decisions for the average consumer.

We took the necessary measures to ensure valid data, but it is possible that different results may have been produced by having the participants choose between two physical pieces of chocolate, rather than selecting one from a comparison line up. By associating a physical decision with the process, a change in behavior may have occurred.

The limited data available for running the analysis process may limit the total significance of the research. One way this can be overcome is through obtaining more information from the participants. A significant way this could be done is by having each participant make a decision for multiple scenarios in order to validate the significance of each individual participant as well as spot significances in different facets of the project. This project was also limited by the informality in which the data was collected. Rather than having randomly chosen a certain set of demographics, the information was gathered and may be heavily skewed toward one or two demographics instead of having a balanced representation across the board of demographics. Improvements for this research could include creating a well outlined verse for an administrator and having physical pieces of chocolate chosen by participants based on the rating and number of comments averaged to obtain the rating. Additionally, it would be good to have done this experiment online, so it could more closely resemble online shopping, rather than doing it in a physical location. By making these changes the project would gain more accurate and beneficial data to better show the significance of the results.

We recommend that business encourage comments and reviews. This could be done through marketing promotions and by making comments and reviews easily available to complete through the placement and clarity in which the process is constructed on company websites. Rather than inhibiting the process, companies could use this as an advantage in the marketing promotion of products by becoming a leader through the relationship they create with the customers through quality products, are received quickly because of good logistics, and who receive great customer service. When customers see these comments on the website, they are able to not only decide between different products within the same product line, but they will have a more informed decision which creates a trust between the customer and the company with the comments. This encourages competition of the market that increases the average quality of products and the standard of living.

References

Adam, C. (2011). Influence of adoption factors and risks on ecommerce and online marketing. *Proceedings Of The International Conference Marketing - From Information To Decision*, *4*21-35.

Bruce, M., Moore, C., & Birtwistle, G. (2004).*International retail marketing*. (1st ed.). Oxford: Elsevier Butterworth-Heinemann.

Godes, D., & Silva, J. C. (2012). Sequential and Temporal Dynamics of Online Opinion. *Marketing Science*, *31*(3), 448-473.

Heydari, A., Teimouri, M., & Jamehshooran, B. (2011). Business ethics and its impact on SCA (Case study Amazon.com). *Interdisciplinary Journal Of Contemporary Research In Business*, *3*(2), 540-555.

Jianwei, L., Peng, H., Hsien-Ming, C., & Lina, Z. (2013). IMPACT OF NATIONAL CULTURE ON ONLINE CONSUMER REVIEW BEHAVIOR. *Global Journal Of*

Business Research (GJBR), *7*(1), 109-115.

Mudambi, S. (2010). WHAT MAKES A HELPFUL ONLINE REVIEW? A STUDY OF CUSTOMER REVIEWS ON AMAZON.COM. *MIS Quarterly*, *34*(1), 185-200.

Sridhar, S., & Srinivasan, R. (2012). Social Influence Effects in Online Product Ratings. *Journal Of Marketing*, *76*(5), 70-88.

Appendix A

Figure 1. Screen shot of Amazon.com's visual product display. This image shows that the first 5 pieces of information received about a product on amazon are image, name, rating, amount of ratings, and price.

Appendix B

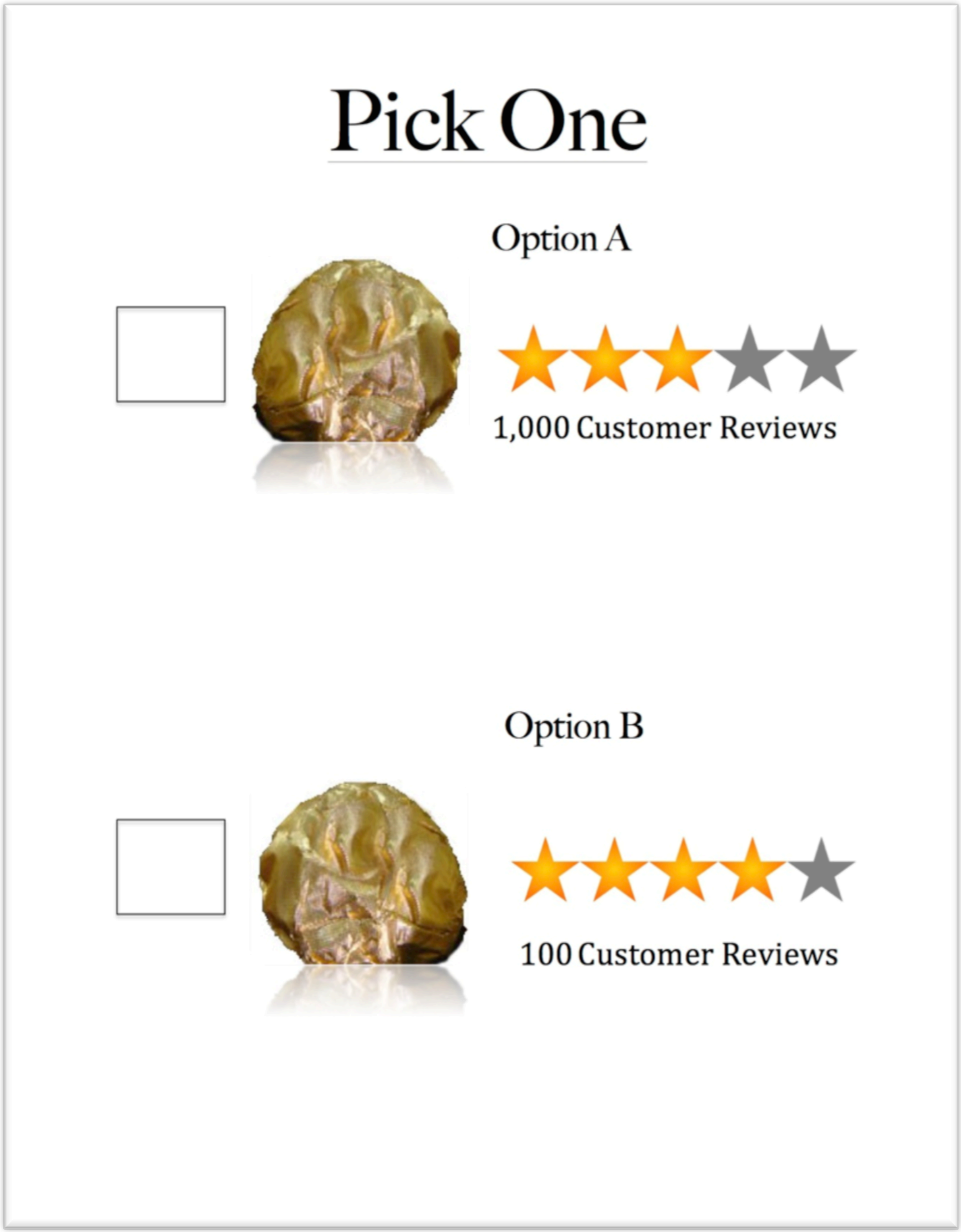

Figure 2. Example sheet subjects marked their choice of chocolate on.

Appendix C

Table 1

Significance level of each independent variable

Variable	Significance
Ratings	.100
Number of Comments	.476
Relationship	.012

Appendix D

Table 2

Binomial Regression Analysis.

	B	S.E.	Wald	df	Sig.	Exp(B)
Rating Difference	-.797	.485	2.703	1	.100	.451
Number of Ratings Difference	-.001	.001	.508	1	.476	.999
Relationship between two variables Difference	-1.833	.730	6.307	1	.012	.160
Constant	-1.908	.961	3.938	1	.047	.148