How Do Consumers Evaluate the Identical Product on Competing Online Retailers?
A Big Data Analysis Approach Using Consumer Reviews

Ji Hye Kang and Sungha Jang
Kansas State University, USA

Keywords: Consumer reviews, big data analysis methods, and online retailers

Purpose of this study: Big data refers to the huge amount of consumer data in a variety of types collected from social, mobile, and local channels. Big data analysis provides unprecedented opportunities for retailers to generate business value by leveraging consumer data analytics. For example, Netflix confidently decided to buy the House of Cards, based on an analysis of the company's subscribers' big data regarding preferred movies, directors, and actors, and made it a big hit (Carr, 2013). That is, the application of big data analysis is already a reality in many industries.

Despite the increasing importance of big data analysis, consumer research in academics is still limited to surveys or experiments, and scant research using big data has been conducted. This motivates our study. According to a ComScore survey (2007), 24% of consumers use online consumer reviews before purchasing a product. Though online reviews can consist of both qualitative assessment (text reviews) and quantitative evaluation (product ratings), many researchers have focused primarily on product ratings only (e.g., Chevalier & Mazlin, 2006; Godes & Mazlin, 2004; Moe & Trusov, 2011) and little has been determined about how consumers utilize other consumers' qualitative assessments.

To address current research's limitation and fully utilize consumer reviews, this study collected both types of consumer review data: a structured form (i.e., review ratings) and an unstructured form (i.e., review text). This study aims to identify 1) consumers' evaluation criteria on a fashion product, 2) positive or negative sentiment toward the product, and 2) the impact of these identified variables on consumers' ratings. For online retailers, Amazon.com and Macys.com were selected for comparison because Amazon.com has recently pushed more its fashion business and its fashion business is now competing with fashion-focused online retailers such as Macys.com (McDowell, 2012).

Methods: Most recently posted consumer reviews on a designer's brand handbag selling at both Amazon.com and Macys.com were collected; a total of 555 reviews (280 review from Amazon.com and 275 reviews from Macys.com) were used for data analysis. Employing text mining analysis, we conducted 1) a *topic modeling analysis* (e.g., categories of product and service attributes that consumers mentioned for the product and service evaluation in their reviews), 2) a *sentiment analysis* of each review (i.e., classification into positive or negative reviews), and 3) a *relationship analysis* between sentiment/topic scores and numeric review ratings. R program and SAS were used for data analysis.

Results and Discussions: *Topic modeling analysis* – After cleaning and organizing the text data (e.g., removing stop words, and stemming words), we found six topics as evaluation criteria: gift, authenticity, roominess, shipping, size, and inside design. Representative words found in consumer reviews for each topic are below.

Table 1. Representative Words for Each Topic.

Topic	Gift	Authenticity	Roominess	Shipping	Size	Inside Design
Representative words	Gift Wife Expect	Authentic Real Love	Pocket Room Hold	Shipped Day Fast	Perfect Size Tote	Inside Compartment Easier

Sentiment analysis – Sentiment of each consumer review was evaluated by counting positive words (e.g., happy) and negative words (e.g., bad). The average sentiment scores of both retailers are positive; 2.996 for Amazon.com and 3.790 for Macys.com. Reviews on Amazon.com ranged from -3 to 15 and those on Macys.com ranged from -2 to 13.

Factors impacting consumers' ratings - In our dataset, 77.8% of Amazon.com's reviewers and 91.6% of Macys.com's reviewers gave a 5-star rating. Due to the skewed responses, we classified the reviews into two responses: 5-star reviews vs. 4 or less star reviews. To examine how the topics and sentiment scores impact on ratings, we set up a logistic regression model for both websites as follows.

$$\frac{P(Rating_i = 5)}{1 - P(Rating_i = 5)} = \alpha + \beta_1 Sentiment_i + \gamma_1 Topic_{1i} + \cdots + \gamma_6 Topic_{6i},$$

where *Sentiment* means sentiment score of review *i*, $Topic_k$ (k=1,...,6) means the probability of that review *i* belongings to Topic *k*. By comparing the coefficients for both websites, we determined whether the effects of topics and sentiment discussed in reviews vary by the online retailers. In the model, the authenticity topic is the base for categorical topic variables.

Table 2. The Impact of Sentiment and Topics on Consumers' Ratings

Parameter	Amazon.com		Macys.com	
	Estimate	S.E.	Estimate	S.E.
Intercept	0.414	0.383	0.788	0.855
Sentiment	0.382***	0.079	0.016	0.109
Gift	5.439***	1.981	2.058*	1.120
Roominess	0.108	0.801	2.301*	1.275
Shipping	-0.800	0.629	1.407	1.261
Size	-0.956	0.643	2.033*	1.172
Inside Design	-0.764	0.695	0.891	1.128

***p<.001, * *p*<.1

Only at Amazon.com, sentiment positively affected review ratings (β_1=0.382, *p*<.001). Macys.com's non-significant relationship between sentiment and rating might be from the fact that Macy's online consumers tend to give 5-star ratings even though they write negative reviews in text. Thus, there is no discriminant effect of sentiment at Macys.com. Regarding the effects of topics, if consumers at Amazon.com are more likely to mention the gift topic compared to authenticity topic, the reviews are more likely to be a 5-star review. The rest of topics have no different effect compared to the authenticity topic. At Macys.com, compared to the authenticity, if consumers mention the gift, roominess, and size topics, the review ratings from those consumers are more likely to be a 5-start review. In conclusion, Macy's online consumers are, in general, more satisfied with the fashion product and have more specific evaluation criteria satisfying themselves. The results suggest that product and service attributes influencing consumers' satisfaction and evaluation are different across online retailer type and its consumers, even on the same product.

References will be provided upon request.