Q&As and Reviews: Substitutes or Complements?

Gaurav Jetley  
Information Systems and Decision Sciences  
Muma College of Business  
University of South Florida  
gauravjetley@mail.usf.edu

Shivendu Shivendu  
Information Systems and Decision Sciences  
Muma College of Business  
University of South Florida  
shivendu@usf.edu

Abstract

Reviews are a major source of post-purchase experiences and have a significant effect on consumer decision making but lack prospective consumer-to-past consumer interaction. QAs facilitates this interaction and provides ability to ask specific pre-purchase questions, and therefore, fills in the informational gaps in the online review systems. However, it’s not clear if the information contained in reviews and QAs is same or different. To answer this key question, we study the complementary and substitute informational content in reviews and QAs for different types of products. We employ a novel method by using measures borrowed from information and measure theory and using topic modeling technique of LDA and find that for high involvement product categories, the reviews and QAs complement each other by providing information on segregated thematic content. For low involvement products, we find that both reviews and QAs contain similar thematic content and might act as substitutes.

Keywords

Reviews, Questions and Answers, Topic Models, Information Theory, Compliments, Substitutes.

Introduction

Reviews have become a major source of information about products and services for customers (Nielson 2012; Zhu and Zhang 2010). Extant literature has shown that reviews have significant effects on consumer decision making process (Chatterjee 2001, Neilson 2013, Dellarocas 2003, Godes and Mayzlin 2014, Ghose et al. 2014) and on vendor sales (Gu et al. 2012, Ghose and Ipeiriotis 2011, Zhu and Zhang 2010). Consumer reviews describe the post-purchase experience of consumers, which provide valuable information about product quality and fit to prospective buyers. In turn, this reduces pre-purchase uncertainty about product valuation (Li and Hitt 2008) and the information and knowledge contained in the stock of reviews also enhances the information transparency and efficiency in markets (Ho et al. 2017).

However, online reviews contain information regarding post-purchase customer experiences and may not address the concerns of prospective customers seeking specific information about product quality, features, or fit. To address this limitation of online customer review systems, large online retailers like Amazon, BestBuy, Target and Walmart have introduced Question and Answers (QAs) sections on their online product pages. Prospective customers can ask specific questions in these sections which can be answered by customers experienced with the product or by the vendor. This interaction between customers and the ability to ask specific pre-purchase questions can also fill in the informational gaps that reviews fail to address (Banerjee et al. 2017). Therefore, QAs mechanism is conceptually quite different from online review

systems. Though both aim to reduce product uncertainty of prospective customers, they might do so by providing information on different aspects of product information.

Using topic modeling, sentiment analysis and other NLP techniques, several studies have investigated the information contained in reviews and user generated content and their role in decision making and helpfulness. Archak et al. (2011) pointed out that the textual content in reviews is important determinant of consumer choice. They extracted product attributes and customer opinions on those attributes by using parts-of-speech tagging approach of Hu and Liu (2004) and syntactic dependency parser to study the economic impacts of these attribute opinions. Ghose and Ipeirotis (2011) studied the different aspects of review text such as subjectivity level features, readability and linguistic correctness and their impacts on sales and perceived usefulness. A recent study compared the thematic content within two sets of human generated information (Huang et al. 2017). The study compared information provided in the textual discussion in analysts’ reports to that of the information provided by managers in the preceding conference calls regarding quarterly earnings by examining the differences in the topic proportions, obtained by LDA algorithm, using Pearson’s chi-square test for homogeneity. They find that analysts add additional thematic information in their reports following the informational content of preceding conference calls through research and show that this additional information is more valuable to investors.

However, previous literature has not investigated the nuances related to two different types of user generated content and their varying informational content. Specifically, in relation to user generated content in online retailer websites, it is not clear if the product information contained in reviews and QAs is same or different and whether this information overlap vary across different types of products, such as search and experience goods and high and low involvement products. We focus on the following questions in this research: What are the key differences in reviews and QAs in terms of the product information that they provide?

If there is a large overlap in information about a specific aspect of the product, such as durability, in both reviews and QAs, then reviews and QAs are substitute mechanisms in reducing the product uncertainty in online markets. On the other hand, if there is little overlap in information about another specific aspect of the product, such as warranty, and this information is contained only is QAs, then these two informational mechanisms (QAs and reviews) are complements.

To answer this key question, we study the complementary and substitute informational content in reviews and QAs for different types of products. We extend the topic modeling for thematic content extraction method used by Huang et al. (2017) by introducing measures and concepts of overlap, divergence and diversity of probability mass functions and random variables from information and measure theory literature to obtain our empirical measures of complementarity and substitutability of information.

Our results suggest that reviews and QAs act as complementary sources of thematic information for high involvement or experience products and as substitute sources of thematic information for low involvement or search products. Specifically, we find that for high involvement product categories, the reviews and QAs complement each other by providing information on segregated thematic content. For low involvement products, we find that both reviews and QAs contain similar thematic content and might act as substitutes.

This study makes several theoretical and practical contributions to literature and has managerial implication. Among others discussed in the Discussion and Conclusions section, the study makes contributions to the online word-of-mouth, text mining and user generated content literature and extends scope for future studies. This study employs a novel text mining technique and provides first evidence of complementarity and substitutability of thematic content in online word-of-mouth, specifically QAs and reviews and how they differ for different types of products. The findings are expected to be generalizable to a large array of online retailers which host a variety of products types or only some product types. The results are also applicable to WOM accumulator websites like Yelp. We make contribution to text mining and topic modeling literature by introducing several measures taken from information and measure theory literature that allow for direct comparison between two or more classes of texts in a corpus in terms of their thematic content.
Relevant Literature

Reviews contain information about different aspects of products which play an important role in consumer decision making and retailer sales. Besides textual reviews, other types of electronic reviews like video reviews, picture reviews and forums in the form of question answers have also made their way to e-markets. This information is very prominently displayed on online product pages to aid customers in decision making. A typical product page consists of product images, description provided by the seller and reviews for the product. E-markets also display concise product information for quick decision making and these can be at the product level (average customer rating, histogram of ratings, etc.) or at the review level (review rating, review helpfulness, review title, etc.). The combination of these serve to assist the customer in making a decision and have shown to make significant impact on product sales (Archak et al. 2011; Cui et al. 2012; Li and Hitt 2008).

Empirical studies have shown that several aspects of reviews like review depth, product type (search & experience), review valence have a significant impact on the helpfulness of the review to customers (Mudambi and Schuff 2010; Yin et al. 2016). Product involvement and valance of the reviews have been significantly associated with product sales (Floyd et al. 2014) as well. Product involvement, defined by Dholakia (2001) as an internal state variable that indicates the amount of arousal, interest or drive evoked by a class of products, can drive the search behavior of customers by increasing the search time for information and the total amount of information gathered (Celsi and Olson 1988; Greenwald and Leavitt 1984). Consumer search behavior aspects that are influenced by involvement include volume of search conducted, active or passive search extent and quantity of information processed (Laurent and Kapferer 1985; Zaichkowsky 1985). Though volume, valence and word counts are important metrics that have been used in previous literate, they don’t capture the entirety of the nuances of the interplay between reviews and QAs (Archak et al. 2011).

Several studies have also employed text mining techniques such as sentiment analysis, topic modeling and other NLP techniques to study the information contained within reviews. These studies show that the positive sentiment in the review title increases the readership of the review and an equal amount of positive and negative sentiment in the review body increases the helpfulness of the review (Salehan and Kim 2016). Archak et al. (2011) extracted product attributes by using POS tagging and extracted customer opinions regarding those attributes by using a syntactic dependency parser to study the economic impacts of these attribute opinions. Ghose and Ipeirotis (2011) studied the different aspects of review text such as subjectivity level features, readability and linguistic correctness in their impacts on sales and perceived usefulness.

A recent addition to the consumer generated information present on product pages are questions and answers (QAs) sections. The presence of reviews and question answers can give the customer greater dimension in capturing essential information to make a decision (Banerjee et al. 2017) and recent research has shown that the volume of not only reviews but also the answered questions can have significant impact on retailer sales for video games (Khern-am-nuai et al. 2017). These QAs can prove to be very helpful in the cases when the product in question has a number of functionalities and performance characteristics that have to be assessed before choosing to buy the product. For example, a customer looking to buy a DSLR camera needs decide whether the body of the DSLR camera is appropriate for their requirements or not. Some of this information can be obtained from the description of camera and customer reviews but more specific information, for example the video quality of the camera compared to another model or lens compatibility from other models can be hard to obtain from the description and reviews alone.

However, previous literature has not investigated the nuances related to two different types of user generated content and their varying informational content. Specifically, in relation to user generated content on online retailer websites, it is not clear if the product information contained in reviews and QAs is same or different and whether this information overlap vary across different types of products, such as search and experience goods and high and low involvement products. This study attempts to explain these nuances and contributes to the literature on online word of mouth and user generated content. We apply a novel text analysis method to study the information content that QAs and reviews provide and the complementarity and substitutability of this information between them for different types of products. We aim to extend this research by studying the differential effects of the complementary and substitute information in reviews and QAs on consumer decision making and retailer sales.
Measurements and Methods

Information & Thematic Content

Following the work of Huang et al. (2017) and Bao & Datta (2014), we quantify our empirical measure of information or thematic content (used interchangeably hereafter) contained in reviews and QAs by extracting meaningful topics from reviews and QAs of several product categories and by using topic modeling method called latent Dirichlet allocation (Blei et al. 2003).

Prior to applying the LDA algorithm to our collection of reviews and QAs, we first assign each product to a category which is homogeneous in product characteristics (i.e. DSLR Cameras, Wine, etc.). The task is to build separate topic models in order to retrieve meaningful topics for each category of products. After the assignment, we run text preprocessing steps like stemming and stop-word removal on the reviews and QAs. Then, we estimate the optimum number of topics for each category of products by using log likelihood and perplexity scores which measure the models’ ability to predict word choices on test data using the model estimated on training data (Blei et al. 2003). We estimate the perplexity score and log likelihood for topic number ranging from 2 to 120 for each topic model built for each category of products. We keep the α and β parameters at 0.1 and 0.01 in accordance to previous literature (Huang et al. 2017; Steyvers and Griffiths 2007). We also manually inspect the validity of the topics obtained from the choices of topic numbers for robustness.

After obtaining the trained topic models for each product category, we subset the outputted Topic-Document matrices (also known as γ-matrix) based on the document being a review or QA. We then use Linear Opinion Pooling to obtain the mixed distributions of topics over reviews of each category of products and do the same for QAs as well. We use these mixed distributions as our measure of thematic content in reviews and QAs for each category of products.

Complementarity & Substitutability

We use measure and information theoretic similarity and dissimilarity measures on the probability mass functions obtained by LDA and LOP which is an extension to the document topic similarity measures used by Steyvers and Griffiths (2007). For measuring substitute thematic content in reviews and QAs, we use Bhattacharyya coefficient (Cha 2007) to approximately quantify the overlap between the thematic content in QAs and Reviews. Bhattacharyya coefficient is defined as $BC(P, Q) = \sum_{i=1}^{n} \sqrt{P(p_i)P(q_i)}$ for discrete probability distributions bounded by $0 \leq BC(P, Q) \leq 1$ with $BC = 0$ when there is full overlap. We also use Hellinger distance: $HE(P, Q) = \sqrt{1 - BC(P, Q)}$ and Jensen-Shannon Divergence: $JSD(P, Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M)$ where D is the KL-Divergence and $M = \frac{1}{2}(P + Q)$ which are other measure of similarity between two distributions and we use them for robustness. For measuring complementarity and substitutability of thematic content on the same scale, we use empirical probability distributions to represent the thematic content in reviews and QAs and then simply measure the probability mass under the curves to get the complementarity and substitutability measures. Here, the sum of complementary and
substitute probability mass for QAs will be 1 and same for reviews. Please refer to Figure 2 for an example of all the measures.

![Figure 2: Example of Complementary and substitute metrics for two distributions](image)

We use entropy from information theory literature as a measure of dispersion of thematic content among reviews and QAs. Specifically, we use Shannon entropy:

\[ H(X) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i) \] 

as a Diversity Index \( D_1 \) or a measure of dispersal. Shannon entropy is higher for flatter probability distributions and lower for peakier distributions. It’s bounded between \( 0 \leq H(X) \leq \log_2(n) \). The idea behind entropy is that an occurrence of an event p which is very likely to occur provides little new information and an occurrence of event q which is a very unlikely to occur provides a lot of information. For example, in the mixed distributions of thematic content in reviews and QAs, if reviews for a certain product only focus on a fraction of thematic content, the total information that reviews provide will be less than if they focused on all the thematic content. This measure will give us a sense of the amount of thematic information contained that reviews or QAs separately with regards to all the thematic information they can cover together.

**Involvement and Product Type**

![Figure 3: Product involvement vs product type (search and experience)](image)

To get our measurements of product involvement, we conducted a paper based survey\(^2\) on 111 graduate students at University of South Florida, Tampa, Florida, to get our estimate of involvement using the same instrument as Lovett et al. (2014) and Ratchford (1987). The average number of valid responses\(^3\) for each

---

\(^2\) USF IRB Protocol Number: Pro00032500

\(^3\) A valid response is when the respondents answer to the question “To what extent are you familiar with the listed product category?” is greater than or equal to 4 on a 5-point Likert scale where 1 is Unfamiliar and 5 is Familiar.
product category was approximately 47 (SD = 9.4). Out of the 29 categories of products, Candy & Chocolate had the lowest involvement level and Laptops had the highest.

We classified the product categories into search and experience by using classification of Nelson (1974) and Laband (1986) in a similar approach used by Lovett et al. (2013) in which two independent judges classified the product categories as search and experience. The intercoder agreement was 72.7% and the differences were resolved by consensus. While we used product involvement as a continuous variable in subsequent analysis, we also dichotomized it into high and low categories using the median split method (Iacobucci et al. 2015) (mean split also resulted in same dichotomization) in order to visual inspection of analysis results. Figure 3 shows a visual representation of distribution of product category involvement by search and experience products where the left figure uses involvement as a continuous variable and the right figure uses the dichotomized version of involvement. This is only for visualization and is not a classification of the products.

**Descriptive Statistics and Results**

To answer the research questions, we gathered reviews and QAs on 355 products belonging to 29 categories of products from Amazon.com in September 2017. For each product, data was collected on its price, age, discount, total number of reviews, total number of answered questions, average star rating given to the product, histogram of review star ratings, all reviews and all answered questions. If the product was unavailable or out of stock, then no price and discount was displayed. A small fraction of the products also did not display the date when the product was first made available. We further created more variables like total word of mouth (WOM) which is just a combination of total reviews and QAs, QA proportion which is the proportion of QAs in total WOM, Volume and percentage of positive (4 & 5 stars), negative (1 & 2 stars) and neutral (3 stars) ratings given to the product, involvement of the product and search/experience classification. The Table 1 below displays the key descriptive statistics of the products in our dataset.

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involvement</td>
<td>355</td>
<td>3.691</td>
<td>0.43</td>
<td>2.977</td>
</tr>
<tr>
<td>Age (Weeks)</td>
<td>331</td>
<td>15,990.16</td>
<td>1,169.16</td>
<td>11,119</td>
</tr>
<tr>
<td>Avg. Rating</td>
<td>355</td>
<td>4.292</td>
<td>0.412</td>
<td>1.888</td>
</tr>
<tr>
<td>Review Volume</td>
<td>355</td>
<td>1,093.15</td>
<td>1,886.74</td>
<td>1</td>
</tr>
<tr>
<td>QA Volume</td>
<td>355</td>
<td>142.515</td>
<td>209.724</td>
<td>1</td>
</tr>
<tr>
<td>Price</td>
<td>341</td>
<td>141.055</td>
<td>303.349</td>
<td>0.95</td>
</tr>
<tr>
<td>Discount</td>
<td>341</td>
<td>5.3</td>
<td>18.224</td>
<td>0</td>
</tr>
<tr>
<td>Positive Rating %</td>
<td>355</td>
<td>0.819</td>
<td>0.17</td>
<td>0.176</td>
</tr>
<tr>
<td>Negative Rating %</td>
<td>355</td>
<td>0.062</td>
<td>0.037</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 1: Descriptive statistics of selected products in dataset**

**Information Complementarity and Substitutability**

| | BC (Q,R) | JSD (Q,R) | HE (Q,R) | H(Q) | H(R) | Emp. Non-Overlap (Q | R) | Emp. Overlap (Q | R) | Emp. Non-Overlap (R | Q) | Emp. Overlap (R | Q) |
|---|---------|-----------|---------|------|------|----------------|----------------|----------------|----------------|----------------|
| BC (Q,R) | -1.00 | -1.00 | 0.31 | 0.08 | -0.66 | 0.66 | -0.47 | 0.47 |
| JSD (Q,R) | -1.00 | -1.00 | -0.28 | -0.09 | 0.67 | -0.67 | 0.47 | -0.47 |
| HE (Q,R) | -1.00 | -1.00 | -0.10 | 0.07 | -0.67 | 0.48 | -0.48 |
| H (Q) | 0.31 | -0.28 | -0.29 | 0.52 | 0.14 | -0.14 | -0.31 | 0.31 |
| H (R) | 0.08 | -0.09 | -0.10 | 0.52 | 0.13 | -0.13 | -0.30 | 0.30 |
| Emp. Non-Overlap (Q | R) | -0.66 | 0.67 | 0.67 | 0.14 | 0.13 | -1.00 | -0.27 | 0.27 |
| Emp. Overlap (Q | R) | 0.66 | -0.67 | -0.67 | -0.14 | -0.13 | 1.00 | 0.27 | -0.27 |
| Emp. Non-Overlap (R | Q) | -0.47 | 0.47 | 0.48 | -0.31 | -0.30 | -0.47 | 0.27 | -1.00 |
| Emp. Overlap (R | Q) | 0.47 | -0.47 | -0.48 | 0.31 | 0.30 | 0.27 | -0.27 | -1.00 |

**Table 2: Correlation matrix of complementary and substitute metrics**

We obtain the distribution of topics in reviews and QAs for each product category. We do this by first sub-setting the document topic matrix for each category of products into two groups of documents. These documents correspond to either reviews or QAs. We then retrieve the mixed distributions for QAs and

---

4 We gathered data on 29 categories of products but only 24 categories were classified into search and experience categories.
reviews for each of these subsets by using linear opinion pooling and obtain the information and measure theory metrics for each category of which a portion is displayed in Table 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wine</td>
<td>0.57</td>
<td>2.29</td>
<td>3.00</td>
<td>0.92</td>
<td>0.08</td>
<td>0.82</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water &amp; Air Filters</td>
<td>0.67</td>
<td>2.66</td>
<td>2.82</td>
<td>0.79</td>
<td>0.21</td>
<td>0.87</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unlocked Cell Phones</td>
<td>0.73</td>
<td>2.94</td>
<td>2.96</td>
<td>0.79</td>
<td>0.21</td>
<td>0.83</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laptops</td>
<td>0.76</td>
<td>3.18</td>
<td>2.72</td>
<td>0.78</td>
<td>0.22</td>
<td>0.82</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSLR Cameras</td>
<td>0.76</td>
<td>3.03</td>
<td>2.49</td>
<td>0.79</td>
<td>0.21</td>
<td>0.79</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glassware &amp; Drinkware</td>
<td>0.78</td>
<td>2.86</td>
<td>3.31</td>
<td>0.78</td>
<td>0.22</td>
<td>0.75</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lamps &amp; Light Fixtures</td>
<td>0.82</td>
<td>2.84</td>
<td>3.08</td>
<td>0.78</td>
<td>0.22</td>
<td>0.71</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headphones</td>
<td>0.82</td>
<td>2.88</td>
<td>2.85</td>
<td>0.82</td>
<td>0.18</td>
<td>0.63</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakeware</td>
<td>0.83</td>
<td>3.07</td>
<td>3.27</td>
<td>0.73</td>
<td>0.27</td>
<td>0.71</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitchen &amp; Table Linens</td>
<td>0.85</td>
<td>3.09</td>
<td>3.30</td>
<td>0.68</td>
<td>0.32</td>
<td>0.74</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Complementary and substitute metrics for 10 out of 29 product categories

Before we study the complementarity and substitutability of reviews and QAs for different categories of product, we investigate the measure and information theoretic metrics of thematic content overlap. Table 2 displays the correlation matrix between the metrics. We can see that BC, JSD and HE are almost perfectly correlated and thus we can use any one of them to gauge the amount of overlap. We also observe that the empirical measures of non-overlapping and overlapping regions are not strongly correlated to BC. This is because the empirical measures are measuring the proportion of only the thematic content in QAs that is also present in reviews and the proportion that is not present in reviews and visa-versa for reviews.

We next conduct visual and simple predictive analysis (using linear and generalized linear models) to gauge the relationship between the thematic overlap measures and product involvement and product type (we do not make any causal statements with the models). To visualize the results, we plot one product from each of the involvement & product type categories with their topic distributions in Figure 4. We then perform tests involving the distributional measures as the predictor variables and involvement and product type as dependent variables displayed in Table 4.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>Involvement (Continuous)</td>
<td>Involvement (Median Split)</td>
<td>Type (Search/Experience)</td>
</tr>
<tr>
<td>GLM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference Class</td>
<td>High Involvement</td>
<td>Experience</td>
<td></td>
</tr>
<tr>
<td>BC (Q,R)</td>
<td>-11.096** (5.239)</td>
<td>102.742** (48.725)</td>
<td>-15.787 (11.762)</td>
</tr>
<tr>
<td>H (Q)</td>
<td>1.496*** (0.422)</td>
<td>-8.691** (3.911)</td>
<td>6.976** (2.761)</td>
</tr>
<tr>
<td>H (R)</td>
<td>-0.829*** (0.295)</td>
<td>5.465** (2.707)</td>
<td>-2.312 (2.175)</td>
</tr>
<tr>
<td>Emp. Non-Overlap (Q</td>
<td>R)</td>
<td>-5.288 (3.848)</td>
<td>59.446* (32.722)</td>
</tr>
<tr>
<td>Emp. Non-Overlap (R</td>
<td>Q)</td>
<td>-6.549 (3.949)</td>
<td>68.929** (34.391)</td>
</tr>
<tr>
<td>Constant</td>
<td>19.579* (9.312)</td>
<td>-169.409** (84.388)</td>
<td>0.677 (2.517)</td>
</tr>
<tr>
<td>Observations</td>
<td>29</td>
<td>29</td>
<td>24</td>
</tr>
<tr>
<td>R Sq. (Adj. R Sq)</td>
<td>0.487 (0.375)</td>
<td>40.899</td>
<td>23.453</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>0.332 (df = 23)</td>
<td>40.899</td>
<td>23.453</td>
</tr>
<tr>
<td>F Statistic</td>
<td>4.367** (df = 5; 23)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Linear and Generalized Linear Models Results

Model 1 and 2, which use involvement as a continuous and dichotomous (median split with high involvement as reference class) dependent variable respectively, show that a lower Bhattacharyya coefficient is positively associated with higher involvement products. Visually, we can observe in Figure 4 that the amount of overlap of thematic content (red) in reviews and QAs for high involvement products, in this case Wine and Laptops, have a much smaller amount of overlap (BC(Q,R) = 0.57 & 0.76) as compared to lower involvement products (BC(Q,R) = 0.82 & 0.85).

Model 1 and 2 also show that higher entropy in QAs and lower entropy in reviews is positively associated with higher involvement products. We can also visually observe that for Wines, a high involvement product, there are certain topics that QAs are concentrated on and reviews though still concentrated on certain...
topics, are a little more diverse. This is in line with the entropy values of 2.29 and 3.0 for QAs and reviews respectively and tells us that for Wines and Laptops, which both belong to high involvement product categories, the reviews and QAs complement each other by providing information on segregated thematic content related to the products. For low involvement products like Kitchen & Table Linens and Lamps & Light Fixtures, Figure 4 shows that both reviews and QAs contain similar thematic content and might act as substitutes. We do not interpret the results of the empirical non-overlap measures as the measures are not significant in both models.

Figure 4: Distributions of topics in QAs and reviews for different categories of products

For Laptops, product category with the highest involvement, we can visually see a stark difference in the diversity of thematic topics that QAs and reviews touch on from that of other categories. Here the QAs are much more diverse ($H(Q) = 3.18$) in the thematic information they provide than reviews ($H(R) = 2.72$) which is the inverse of what we see in the lower involvement products. These results suggest that for laptops, the QAs add value by displaying text containing a diverse range of thematic information and exhibit a complementary behavior to reviews, while reviews mostly focus on a few.

Model 3 uses product type (search or experience) as the dependent variable with experience as the reference class. The results show that higher entropy in QAs and lower empirical non-overlapping region in QAs given reviews is positively associated with search type products. We can also visually see this in Figure 4 where Laptops and Kitchen & Table Linens have lower empirical non-overlap in QAs given reviews (0.68 & 0.78) than Wine and Lamps & Light Fixtures (0.78 & 0.92). This suggests that QAs start to discuss different thematic content from reviews when we examine experience products (as compared to search products) and thus show a complementary behavior to reviews in experience products.
Discussion & Conclusion

In this research, we study a new mechanism in online marketplaces, namely QAs. This mechanism provides a way for interaction between the prospective customers who ask pre-purchase questions to lower their fit and quality uncertainty and former customers who provide post-purchase answers. This interaction between customers and the ability to ask specific pre-purchase questions can also fill in the informational gaps that reviews fail to address (Banerjee et al. 2017). We empirically study this mechanism in conjunction with customer reviews which have been shown to have significant effects on consumer decision making and retailer sales. We study QAs, a mechanism which is conceptually quite different from online review systems though both aim to reduce product uncertainty of prospective customers, to analyze the information that they provide to consumers. Specifically, we try to find if the information contained in reviews and QAs is complementary or substitute for different types of products.

Our results suggest that reviews and QAs act as complementary sources of thematic information for high involvement or experience products and as substitute sources of thematic information for low involvement or search products. Specifically, we find that for high involvement product categories, the reviews and QAs complement each other by providing information on segregated thematic content. For low involvement products, we find that both reviews and QAs contain similar thematic content and might act as substitutes. We also find that QAs start to discuss different thematic content from reviews when we examine experience products (as compared to search products) and thus act as complementary sources of thematic content to reviews.

This study makes several theoretical and practical contributions. First, the study makes contributions to text mining and user generated content literature and extends scope for future studies. Using novel text mining techniques, we provide first evidence of complementarity and substitutability of thematic content in online word-of-mouth, specifically QAs and reviews and how they differ for different types of products. This can also potentially explain consumer search behavior online for different types of products. Second, the results highlight the importance of product involvement and product type in WOM content generation by the consumers. Third, the study brings several managerial implications for online retailers and sellers. The method we use can be used by the retailers to tailor and optimize product pages that display the WOM (QAs and reviews) information more prominently to consumers for products that have high informational needs. More information for certain products can reduce fit uncertainty and influence consumer decisions and in turn product sales. The online sellers can also potentially increase the customer experience by focusing on answering unanswered questions or correcting wrongly answered questions for certain types of products where QAs are complimentary source of information. Fourth, the findings are expected to be generalizable to a large array of online retailers which host a variety of products types or only some product types. The results are also applicable to WOM accumulator websites like Yelp. Fifth, we make contribution to text mining and topic modeling literature by introducing several measures taken from information and measure theory literature that allow for direct comparison between two or more classes of texts in a corpus in terms of their thematic content. One of several possible avenues for future work will be to explore the effects of WOM thematic content substitutability and complementarity for individual products of different types on consumer decision making and retailer sales.

Acknowledgements

We greatly appreciate the reviewers’ suggestions and feedback. We also thank Aravind Reddy Meenige for lending his expertise in web scraping and data wrangling for gathering data for the research.

REFERENCES


Mining Text and Reviewer Characteristics," IEEE Transactions on Knowledge and Data 
Engineering (23:10), pp. 1498-1512.


Huang, A. H., Lehavy, R., Zang, A. Y., and Zheng, R. 2017. "Analyst Information Discovery and 

Nuanced Understanding of the Statistical Properties of a Median Split," Journal of Consumer 

Sales: The Case of Amazon Answer.

Statistics, pp. 517-521.

research), pp. 41-53.


research (50:4), pp. 427-444.


consumers-trust-in-earned-advertising-grows.html.

Nielson. 2013. "The Reviews Are In: Yelp Users Are Four-Star Consumers." Retrieved February 1, 2018, 
four-star-consumers.html.


Reviews? An Empirical Investigation of Confirmation Bias in Online Word of Mouth," Information 
Systems Research (27:1), pp. 131-144.

341-352.

Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of 