

DO ONLY REVIEW CHARACTERISTICS AFFECT CONSUMERS' ONLINE BEHAVIORS? A STUDY OF RELATIONSHIP BETWEEN REVIEWS

Guofang Nan
College of Management and Economics
Tianjin University,
Tianjin 300072, China
gfnan@tju.edu.cn

Jiaorong Yang
College of Management and Economics
Tianjin University,
Tianjin 300072, China
18801900974@163.com

Runliang Dou*
College of Management and Economics
Tianjin University,
Tianjin 300072, China
drl@tju.edu.cn

ABSTRACT

With the limited screen size of mobile phones, consumers cannot read product reviews as freely as they do at computer terminals. To rectify the insufficient number of helpful review voters, consideration should be paid not only to the value reviews provided by consumers but also to the relationship between reviews so as to recommend the optimal reading order for consumers and enable them to obtain more product information. In this paper, we measure product uncertainty to show the relationship between reviews with consumers continually reading reviews and to explore the influence of their perception change on their online behaviors. The product uncertainty is computed by the improved Shannon entropy, which measures the product review text on websites. With our data collected from Amazon.com by python coding, three significant findings are detected as follows. First, regardless of the order the reviews are sorted in, whether by the most recent or the most helpful, the results demonstrate that the varied product uncertainty of a review has a significant relationship with its helpful voters. Moreover, the more the product uncertainty is varied by a review, the more possible it will be for the product to obtain voters. Second, when analyzing the influence of product uncertainty on consumers' purchasing behavior, we find that the lower the product uncertainty computed by many reviews, the higher the product rank, and the more likely consumers are to purchase the product. Third, by exploring the experimental results with knowledge of relevant behavior psychology, we offer different meanings for display reviews on different e-commerce sites.

Keywords: Product uncertainty; Online consumer reviews; Consumers' behaviors; Shannon entropy

1. Introduction

With the widespread adoption of mobile phones, mobile commerce has become increasingly popular. In particular, Alibaba, Jingdong, and Amazon will strategically upgrade m-commerce. The Nilson Report states that more and more Chinese people have begun to buy goods and services through mobile devices, particularly those in 3rd and 4th tier cities. With a low price, small volume, and powerful functionality, smart phone industry gains breakthroughs. In China, smart phone penetration among those who are less than 30 years old was 62% by 2014; computer penetration is only 35%. In contrast to other platforms, mobile phones have two distinct characteristics [Shankar & Balasubramanian 2009; Clarke 2001]. First, mobile phones have real-time functionality and high flexibility as they can be used anywhere and at any time due to their small size and powerful functions. Second, using mobile phones could enhance personalized service and, in particular, improve the interaction between retailers and

* Corresponding author

consumers and the repeat behavior after the geographical location is considered. However, the screen size of a smart phone is usually 3.5-5 inches; for example, the screen size of iPhone 6 is only 4.7 inches. The different screen size causes users to behave differently. With the increasing cost for consumers to search and understand products by using mobile phones, it is urgent to guide them to efficient e-shopping by offering more shopping references, which is not only a problem of the mobile phone user interface but also an optimization problem of the whole shopping procedure, including the recommendation searching system and the recommending system.

For instance, Figures 1 and 2 reflect the users' reviews displayed on mobile phones and computers, respectively. Figure 1 shows that there are only 2 review comments displayed on iPhone 6s, one of which cannot be fully displayed, while Figure 2 shows that there are 5 review comments fully displayed on a computer screen. Therefore, it is impossible for consumers to read thousands of reviews with a 4.7-inch mobile phone.

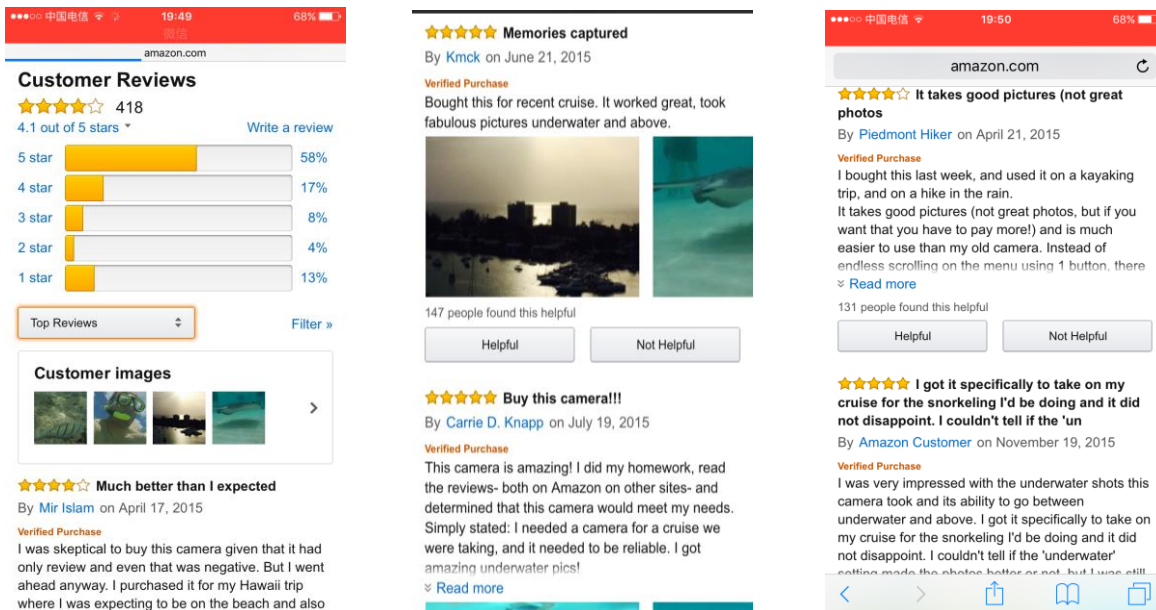


Figure 1: Reviews Are Shown on a Mobile Phone

Traditionally, the product reviews are ordered on the basis of helpful votes, from the most helpful to the least helpful on a traditional e-commerce platform. Consequently, not all the reviews could be efficiently voted on. In this case, the traditional recommending systems do not work effectively. However, if the reviews on an e-commerce platform are ordered by review time from most recent to least recent, certain useless reviews would restrain consumers from buying the product, which is more significant when a consumer uses a mobile phone to shop.

Most literature has studied how product sales are influenced by product reviews and which factors of the reviews constructively contribute to helpful votes. Nevertheless, the review is treated as an independent entity, and interactions between different reviews are ignored in most previous studies. For instance, certain researchers apply characters from an independent review, such as readability, objectivity, and reviewer's reputation to explain why it receives helpful votes. However, the researchers neglect the influence caused by the relationship between reviews. Assuming an extreme case, there are two similar reviews describing the same attribute of a product with the same sentiment. Intuitively, the first review that ranks higher will gain more helpful votes, since consumers have previously acquired product information from it, which has provided more valuable information. Consequently, the second review only receives fewer helpful votes or no helpful vote. Conversely, the second review will gain more helpful votes if the first one is not read by consumers. Similarly, the previous studies on the influence of the reviews on product sales address the reviews independently, rather than analyze the logical relationship of the reviews from their semantic and perceptive aspects. For instance, consumers' perception of a product after reading ten reviews could not be equal to the sum perception of reading them separately due to the logical relationship between reviews. Therefore, this paper's objective is to study product reviews from this new perspective and explore how the relationship between reviews impacts consumer behavior.

★☆☆☆☆ Disappointing
By [crowdance](#) on 17 January 2016
Colour Name: Black | Style Name: Camera | **Verified Purchase**
Disappointed with this purchase. The quality of the pictures is a bit ropey and it doesn't perform well in low or bright light. I have only had it for two weeks, but having tried it indoors and out, most of the time I am leaving it at home. I had a sony compact a few years ago and it was really good, so i was expecting much more of this product. My advice, spend a bit more and get a better product.
▶ [Comment](#) | 3 people found this helpful. Was this review helpful to you? [Report abuse](#)

★★★★★ Stylish, lightweight, suitable for non-tech-savy people
By [HRB](#) on 11 May 2016
Colour Name: Black | Style Name: Camera | **Verified Purchase**
I bought this as a birthday gift for my mother, who was previously using her iPad to take photographs (much to the annoyance of fellow human beings across the planet). She was becoming increasingly irate with the poor quality of her iPad photographs and complaining about the mysterious red line that appeared on every single photo (this was her red leather iPad cover partially obscuring the lens). Anyway, all iPad based woes are now long forgotten as she has this handy little camera instead. She required assistance to set the time, date and language though my mother requires assistance in telling the difference between a telephone and a TV remote so I wouldn't worry too much about that. The photo quality is good, it's stylish, lightweight, sturdy and easy to use.
▶ [Comment](#) | 2 people found this helpful. Was this review helpful to you? [Report abuse](#)





★☆☆☆☆ Poor photo quality
By [Nicki](#) on 2 January 2016
Colour Name: Black | Style Name: Camera | **Verified Purchase**
Purchased this to replace my old 9 megapixel camera. I set it up and took a few photos and they looked blurry on screen. I plugged it into my laptop thinking that the screen wasn't a high enough resolution to properly view the pictures only to discover that they were in fact terrible quality and some were blurred. I altered the settings and tried a few different types of lighting to check that it wasn't just a one off and unfortunately was disappointed to discover that it made no difference, and my old 9 mp camera provided better results. We had no choice but to return the camera and will be looking for an alternative.
▶ [Comment](#) | 9 people found this helpful. Was this review helpful to you? [Report abuse](#)

★★★★★ Five Stars
By [hugo couto](#) on 2 December 2014
Colour Name: Silver | Style Name: Camera | **Verified Purchase**
love it really good service good camera good quality and quick delivery recommend
▶ [Comment](#) | 24 people found this helpful. Was this review helpful to you? [Report abuse](#)

★☆☆☆☆ Very disappointing
By [Tim](#) on 10 August 2015
Colour Name: Black | Style Name: Camera

What do you want to know about Sony DSCW800 Digital Compact Camera (20.1 MP, 5x Optical Zoom) - Black?
[See all 238 answered questions](#)

Customers also viewed these items

-  **Sony DSCW630 Digital Compa Camera - Black (20.1MP, 8x Optical Zoom) 2.7 inch LCD**
by Sony
£79.00
★★★★☆ 57
-  **NIKON VNA871E1 NIK2107 - (Cameras > Digital Cameras)**
by Nikon
£63.99
★★★★☆ 75
-  **Canon IXUS 160 Point and Sho Digital Camera - Black**
by Canon
£72.10
★★★★☆ 49
-  **Sony DSCW220 Digital Compact Camera with Wi-Fi ar NFC (18.2 MP, 10x Optical Zoom) -...**
by Sony
£139.99 - £159.99
★★★★☆ 58

Send us feedback
How can we make Amazon Customer Reviews better for you?

Figure 2: Reviews Are Shown on a Computer Screen

Specifically, whether consumers vote for the helpfulness of a review depends on how much they know about a product. For a consumer, the greater the perception obtained from a review, the more likely he/she would be to consider it helpful. As the consumer's perception has changed after reading that review, he/she will decide whether to vote for the product with his/her new perception. To an extent, the finding could be explained by the concept of a reference point from prospect theory, which has been proven as a notable factor that affects decisions in behavioral psychology. In this situation, the reference point changes constantly with continuous reading. In other literature, consumers' perception is replaced by their uncertainty for a product that they are familiar with. In sum, a consumer updates his/her perception- of a product by reading product reviews, and his/her behavior (purchasing or voting) demonstrates that their perception has changed. Therefore, product uncertainty acts as not only a factor to reflect relationship between reviews but also a measurable variable to influence consumers' online behaviors in the context of continuously reading.

Numerous prior studies have investigated product uncertainty by questionnaire, which may suffer from limited input and conclusions reached based on limited data. To simplify this difficult task, some researchers merely use review volume (the number of reviews) or value (ratings) as substitutes to express the review information. Only a few researchers analyze the reviews textually and semantically. It is an expedient strategy to use numeric data representing review content for the complexity of free-text's structure and grammar. However, the advent of text analysis tools, such as NLTK, enables us to address the textual content automatically and effectively. Hu et al. [2014] note that consumers make purchasing decisions not only based on the review rating but also on the partial information of the review text, which cannot be replaced by the rating noted. Knight [1921] defines uncertainty as ignorance or partial knowledge of a product. In recent years, researchers enrich the concept of uncertainty in the context of electronic commerce. We only focus on the product uncertainty from a textual review. For example, the reviews on the attribute of one camera on Amazon.com are self-conflicting, such as "the lens is wonderful" and "the lens is somewhat blurry". The conflict will increase consumer confusion about the camera lens and their purchasing decision. To some extent, the external hints, such as reviewers' reputation, could help consumers in judging which review is more convincing. However, one product will easily obtain thousands of reviews, some of which positively state one attribute of the product, while others negatively present the same attribute. If they evaluate one product by relying on the external hint of the reviews, it will be easy for a potential consumer to feel exhausted and bored instead of funny and relaxing in a real shopping experience. Furthermore, the product uncertainty will be influenced by consumers' personal preferences in certain research. If the uncertainty is objectively measured only from the aspect of textual reviews, consumers' personal preferences will be ignored. However, with the consumers' preferences, the personal product uncertainty could be re-computed by adjusting the product attribute weight.

In this paper, considering that mobile phones have small screens, we intend to find an approach that could enable the reading of the whole review and that could reflect relationship between them. We improve the Shannon entropy, which is a typical approach to compute uncertainty from review textual content, to understand consumers' online behaviors. On this basis, we raise our research questions as follows:

Q1. Does product uncertainty varied by a review significantly contribute to its helpful votes?

Q2. Does product uncertainty play a significant role in product rank?

How much product uncertainty varied by a review, to some extent, can reflect how much a consumer's perception updated after reading it? Specifically, the change in product uncertainty by reading a review drives consumers to vote for the review, rather than for the information it contains. Moreover, the measured uncertainty by the same review tends to vary according to its position (all reviews could be ordered by most helpful or most recent in Amazon).

Our work supplements prior research in several aspects. First, previous studies about reviews explore the relationship between reviews and consumers' online behaviors based on the simple statistics of product attributes and their emotion, while the intrinsic mechanisms remain undetermined. In fact, the existing approaches tend to treat many reviews as independent entities, ignoring the relationship between them to some extent. In this paper, we focus on both the characteristics of an individual review by text mining and the relationship between reviews to measure product uncertainty. Second, a questionnaire is the typical approach used to evaluate product uncertainty in previous research. We enrich this field via computing product uncertainty from reviews' textual content by the improved Shannon entropy. Furthermore, this paper also provides two practice contributions. On the one hand, product manufacturers could better understand their product at the attribute level and launch specific marketing activities, such as providing more convincing descriptions for the confusing attributes in product reviews. On the other hand, platforms, such as Amazon.com, could apply product uncertainty as a reference factor to rank or recommend reviews, thereby guiding consumer shopping and improving product sales.

To summarize, our paper aims to understand how the product uncertainty of contextual reviews affects consumers' behaviors based on the concept of a continuous reading process. This paper could be regarded as an extension of the prior work of Archak et al. [2011]. We also enrich the concept of sequential belief updating from consumer reviews by product uncertainty and conduct empirical experiments to demonstrate this with real data from Amazon.com.

The remainder of this paper is organized as follows. In Section 2, we provide a brief literature overview. In Section 3, we elaborate on our methods for the computational process of product uncertainty. The models are represented in Section 4. In Section 5, we show our research results. Discussions are provided in Section 6. Finally, we discuss managerial implication and conclusions.

2. Literature Review

Several key concepts and certain basic textual processing methods are reviewed in this section, including a brief summary of the research findings in the mobile shopping field, the perceptive product uncertainty, the method of product attributes identification and the emotion extraction to cope with the complex structure and grammar of the text.

2.1. Mobile shopping

Online shopping by mobile phones has made our daily life more convenient and comfortable [Kiang et al. 2000; Kiang & Chi 2001]. Chong [2013] investigates the determinants of m-commerce usage activities and shows that age and gender have no significant impact on mobile shopping; in addition, high educational levels are more likely to use m-commerce. Chong et al. [2012] compare the m-commerce adoption decisions in China and Malaysia and demonstrate that trust and social influence have significant and positive relations in both countries. Martín et al. [2012] note that consumers' value for the firm is the key element to determine the use of mobile commerce; firms could gain enough knowledge of their consumers and make personalized recommendations by e-commerce. Yang [2010] says that Utilitarian performance expectancy and Hedonic performance expectancy are positively related to the attitude toward using mobile shopping services. Yang & Kim [2012] believe that mobile shopping could help consumers obtain more information when they are on the move; it also saves time in finding a store and products. Moreover, mobile coupons enable consumers to purchase products at low prices. Customers can explore a mobile site in a real store to obtain reviews and experience products at the same time. Wong et al. [2012] verifies that perceived usefulness and perceived ease of use has a positive significant relationship with m-shopping adoption. By using a questionnaire, Wang et al. [2015] find that m-shopping customers tend to purchase products repeatedly. Huang et al. [2016] find the frequency of purchasing on the web will decrease after providing a mobile channel.

Compared to the current questionnaire method, we intend to investigate the relationship between online reviews and consumer behavior by text analyzing and quantitative analysis; unstructured data processing will be the major challenge.

2.2. Online product uncertainty

The primary literature on the uncertainty definition originated from Knight [1921] who describes uncertainty as the partial knowledge of a product. Recent studies associate uncertainty with risk and information asymmetry [Hong & Pavlou 2014; Dimoka et al. 2012; Li & Hitt 2008; Zhang et al. 2014]. In the context of e-commerce, product uncertainty is the main obstacle restricting consumers from purchasing. In contrast to a physical store, consumers could touch, observe and experience the product, talk to the sellers face-to-face to decrease the uncertainty and avoid the risk of purchasing an improper product, or professionally solve the product quality problem, conatural deficiency and other troubles. The online product uncertainty can be divided into two categories. One category is seller uncertainty, which reflects that buyers do not trust sellers; this usually occurs on a C2C platform. The other is product uncertainty, which is caused by the limited or asymmetrical information of products, particularly when various products are not described with standard features. For simplification, our paper focuses on the latter.

Consumers gradually reduce product uncertainty at online stores by referring to the textual content of the product description and reviews. Recent studies provide meaningful results. For example, Hong & Pavlou [2014] state that a product sale is negatively related to product uncertainty by conducting a questionnaire survey; this conclusion is also shown in Dimoka et al. [2012].

Despite consumers' personal preferences and heterogeneity having a significant influence on product uncertainty, different consumers have different understandings of the same content and would allocate different weights to the same product. For instance, relying on consumers' behavior recorded by the platform, the website understands that a consumer focuses more on a particular attribute of a product. Hence, the weight of this attribute could be amplified when measuring the product uncertainty for him/her.

2.3. Product reviews content

The significant influence of product rating on a consumer's purchasing decision has been verified in many studies [Hu et al. 2016; Kolomiiets et al. 2016]. Katz & Elihu [1955] find that the word-of-mouth effect is a key factor in influencing consumers' purchasing decision, which is 7 times larger than that of an e-magazine and 2 times than that of an advertisement. Liu et al. [2008] demonstrate that coding review content by manpower is an extremely tedious task. However, with the development of text analysis tools, this coding helps us gain a better understanding of consumers' behaviors by decoding textual content automatically. Hu et al. [2014] conclude that favorable reviews could boost product sales by observing the panel data of DVDs, books and videos from Amazon.com. Similarly, by examining the impact of reviews on hotel room sales, Ye et al. [2009] obtain the conclusion that high ratings can significantly increase the number of bookings, while the variance of ratings has an opposite effect. However, in the film market, Duan et al. [2008] find a slightly different result in that WOM valence (ratings) does not have a direct effect on revenue, while volume does.

As the number of reviews increases, the information overload problem becomes more obvious. This problem has prevented potential consumers from fully utilizing the reviews; that is, it is impossible for consumers to read all the reviews. To improve the consumers' purchasing efficiency, online review systems, such as Amazon.com, provide a voting mechanism for consumers to express their attitude for a review, and then the website displays these reviews based on the number of helpful votes. Certain studies attempt to address such issues. For instance, Martin & Pu [2014] propose a helpful votes' model to select reviews containing valuable information from the viewpoint of meta-data, semantic and emotion. Kim et al. [2006] regards structural, lexical, syntactic, semantic, and meta-data as features of reviews to build a framework of content analysis. Forman et al. [2008] adopts the reviewers' reputation, the readability of the review text and the subjectivity of the review text to conduct analysis. Liu et al. [2008] consider reviewers' expertise, writing style and timeliness of the reviews as features. All the aforementioned papers conduct empirical studies to support their models and, to some extent, solve these issues.

While previous studies have explored the factors for helpful votes by decomposing the review content, to the best of our knowledge, few papers truly investigate the internal mechanism between reviews and consumers' behaviors, particularly at the product attribute level.

2.4. Product attributes identification

It is a challenge for consumers to manually convert reviews' textual content to a fine-gained level with multifarious product attributes and their corresponding emotions. Fortunately, this effort could be realized with text-mining tools. Supervised and unsupervised approaches are the two branches used to extract product attributes [Xu et al. 2011]. The supervised approach applies machine learning methods, such as SVM and Naive Bayes, while the unsupervised approach usually relies on lexicon.

For product attribute identification, the typical approaches abstract the frequency occurrence of nouns [Netzer et al. 2012; Bafna & Toshniwal 2013]. To overcome this complicated linguistic issue, Putthividhya & Hu [2011] regard abstracting attributes as an entity recognition problem; Ganu et al. [2013] apply Bays to identify attributes. However, the match results will not be precise if certain key nouns are replaced by noun phrases. Regarding the product

attributes' emotion identification, each emotion could be expressed by varied adjective words, and all of them convey information to consumers [Yu et al. 2011]. Therefore, due to the language's complexity, we considered all relevant adjectives to identify consumers' emotion in this paper.

To match the noun and its corresponding adjective, one alternative approach is to compute the distance between a noun and its closest adjective to pair them. Another well-accepted approach applies PMI (pointwise mutual information) to calculate the match degree [Quan & Ren 2014]. Putthividhya & Hu [2011] suggest that one noun may be adorned by several adjectives and applies $\text{window} \pm 2$ to match them, which means that, if a noun is located first, its adjectives can be found within the distance of two words. Based on the above conclusions, we decide to utilize the third approach to obtain pairs with an improvement of setting a limitation by punctuations. That is, if a noun and an adjective belong to different sentences, the pair could not be combined, including when the condition is $\text{window} \pm 2$.

3. Research Methods

Our objective is to explore the relationship between product uncertainty and consumers' online behaviors. In the following, we introduce the technique approaches to measure product uncertainty.

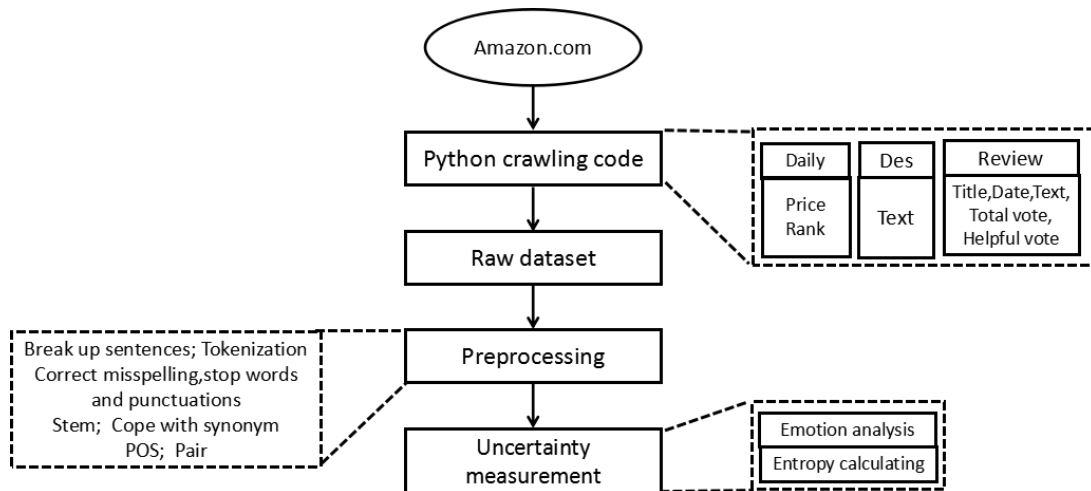


Figure 3: The Mining Process

Figure 3 presents the system of text mining and the measurement process of product uncertainty. We collect a dataset from amazon.com by using python crawling code, and more details can be referred to Section 3.1. The text pre-processing methods are then applied to address language's complexity problems in Section 3.2. Finally, details on how to measure product uncertainty are provided in 3.3.

3.1. Data collection

We collected our data from Amazon.com, which is the leader in electronic commerce, by using python coding. Specifically, we selected ten products with several hundred reviews, half of which are digital cameras presented as typical search goods, and the remaining are mobile applications considered as experience goods [Chen & Tseng 2011; Mudambi & Schuff 2010]. For each product, we also collect the product description, the review content and the review textual content, and we review meta data such as posted date, helpful votes, total votes, and ratings from the detailed product website. Additionally, daily data on each product's rank, price, the average ratings, and the number of reviews were gathered during the 8 weeks from April 10, 2015 to June 4, 2015.

The summary statistics are provided in Table 1. Each product has enough reviews for us to analyze, and there are a total of 2309 and 3475 reviews for search goods and experience goods, respectively. It is obvious that the average prices of search goods are much higher than those of experience goods.

Table 1: Summary Statistics of a Full Sample from Amazon.com

Type	Product	Reviews_number ¹	Avg_rank ²	Avg_price ³
Search Goods (Cameras)	Canon	867	466	161
	Nikon	238	121	225
	Sony	386	132	676
	Olympus	302	1701	262
	Fujifilm	516	68	250
Experience Goods (Apps)	SoundCloud	981	178	0
	Tube+	950	205	0
	King of Thieves	464	1957	0.99
	Bank Escape Pro	595	3375	2.99
	Super Stickman Golf	485	454	0

3.2. Text pre-processing method

To process the noisy and complex review textual content and convert them into structured data, we adopt several typical approaches presented in the previous studies. The main pre-processing steps can be observed in the left of Figure 3; more details are introduced as follows:

- ◆ Breaking up sentences. Segregate sentences by sentence punctuations to distinguish the scope of a sentence.
- ◆ Tokenization. Utilize tools packages in python NLTK to separate sentences into single words.
- ◆ Correcting misspelling, removing stop words and punctuations. Use the epython pattern package to correct the misspelled words and to eliminate stop words according to lexicon that exist in NLTK packages.
- ◆ Stemming. Convert word to its root form, such as change the word cars to car. Similarly, nouns with the same meaning but diverse forms could be treated in the same way.
- ◆ Coping with synonyms. Using the word net dictionary, we unify the synonyms, particularly adjectives, to one basic word to effectively reduce the analysis dimensions.
- ◆ POS and Identification. Annotate each word with its part of speech and mark the word with noun, verb, or adjective, in that product attributes are usually expressed as nouns and emotion words as adjectives.
- ◆ Pairing. We apply a window of two words to match attributes and their corresponding emotional words. In particular, we first record the location of the noun, which is previously identified as a product attribute; we then search for its window ± 2 adjective words that refer to the noun until meeting the sentence pause. Thus, we could obtain many useful pairs, based on which the information entropy is calculated.

3.3. Product uncertainty measurement

The pairs of product attributes and their corresponding emotional words are obtained following the above steps. We adopt a widely accepted approach, the Shannon information entropy, to measure uncertainty. Introduced in early 1948 [Shannon 1948], the Shannon information entropy is widely employed in academic studies. Zhang & Tran [2009] utilize this approach to identify important terms from reviews according to helpful votes. Similarly, Montemurro & Zanette [2002] employ the Shannon entropy to judge the linguistic role of each word.

In this paper, we apply the Shannon information entropy to measure the product uncertainty. Intuitively, if the number of positive words equals that of negative words for a product's attribute in reviews, the potential consumers may be confronted with a dilemma. According to the Shannon entropy, the closer occurrence of each condition is, the larger the uncertainty of the event is. Let $S = \{s_1, s_2, \dots, s_n\}$ be a set of product attributes without repetition, where s_i means a product attribute, and $R = \{r_1, r_2, \dots, r_n\}$ is a set of all reviews that are ordered by a certain rule (e.g., most helpful or most recent), where r_i means one of the reviews. Let $D_i = \sum r_i$ be a reviews collection that contains all the reviews with ranking before r_i . In addition, we define

$$P_{pos_{it}} = \frac{pos_{it} + 1}{pos_{it} + neg_{it} + 1}, \quad (1)$$

¹ The number of reviews

² The average rank of the product during eight weeks

³ The average price of the product during eight weeks

$$P_{neg_{it}} = \frac{neg_{it} + 1}{pos_{it} + neg_{it} + 1}, \tag{2}$$

where

- n The number of product attributes (specifically, in accordance with previous studies, we only apply product attributes with the top 30% frequent occurrence of noun words in reviews)
- m The total number of reviews
- pos_{it} The times that the product attribute s_t is described to be positive words in D_i
- neg_{it} The times that the product attribute s_t is described to be negative words in D_i
- $P_{pos_{it}}$ Positive emotional rating of product attributes s_t in D_i
- $P_{neg_{it}}$ Negative emotional rating of product attributes s_t in D_i

When a product attribute only reflects positive emotion and no one writes the negative aspect in D_i , a mathematical problem will arise and result in infinite entropy. To avoid this issue, we modify the typical calculation of the occurrence equation as Eq. (1) and Eq. (2), assuming that both one positive emotion and one negative emotion exist for every product attribute. In particular, when consumers read a product attribute and its emotion, it appears impossible for them to trust it immediately; thus, the attribute uncertainty changes constantly in the process of reading reviews. In other words, when the Shannon information entropy is used to calculate product uncertainty, we set a basic condition that one positive word and one negative word currently exist for the product attribute. It is noted that, if there are only a few reviews, the importance of product uncertainty is very sensitive, and the sensitivity of the product uncertainty declines with an increase in corresponding reviews. Thus, the Shannon entropy is effective for measuring the product uncertainty for a product that has been reviewed many times.

Then, for product attribute s_t in D_i , the uncertainty is calculated as follows:

$$H_{it} = -(p_{pos_{it}} \times \ln p_{pos_{it}} + p_{neg_{it}} \times \ln p_{neg_{it}}). \tag{3}$$

Hence, considering all the product attributes, the total average product uncertainty in D_i is:

$$AH_i = \frac{1}{n} \sum_{t=1}^n H_{it} \tag{4}$$

Varied uncertainty by one review r_i could be measured by the following equation:

$$\Delta H_i = AH_i - AH_{i-1} \tag{5}$$

where AH_i denotes all the product attributes in D_i , and ΔH_i refers to varied uncertainty by review r_i .

In fact, consumers obtain the overall perception of a product when they read the product description, and their perception is updated as they read an increasing number of product descriptions. Hence, when varied uncertainty by reviews is measured, the influence of the product description should be considered. From the technical methods' perspective, the uncertainty could be resolved by regarding product description as the first review to compute. Moreover, as we have repeatedly previously stressed, it is the change of the utility that makes consumers decide to vote for a review. Thus, it appears more objective to reflect the true value of the review by calculating the varied uncertainty by review, rather than by calculating the entropy of the review content. For example, a product description and its four reviews are defined as Des, r_1, r_2, r_3, r_4 , respectively. We intend to compute the varied uncertainty by r_3 . According to Eq. (4), the total average product uncertainty is first measured by Des, r_1, r_2, r_3 ; then the total average product uncertainty is measured by Des, r_1, r_2 only. The difference ($\Delta H_3 = AH_3 - AH_2$) is the varied uncertainty, measured by r_3 and the true value of r_3 . In particular, the utility is generated by the difference between the current situation and the reference point rather than by the current situation only.

Amazon.com provides two methods to sort reviews: on the order of most helpful and most recent. As review position affects the product uncertainty significantly, both situations are considered in our paper. Figure 4 depicts the product uncertainty of search goods under the condition that all reviews are ordered by most helpful according to Eq. (5), and Figure 5 depicts the condition under which all reviews are ordered by most recent. Figure 6 and Figure 7

depict product uncertainty of experience goods under the condition that all reviews are ordered by most helpful and all reviews are ordered by most recent, respectively. The horizontal axis denotes the number of reviews in the certain rank condition, and the vertical axis denotes the product uncertainty.

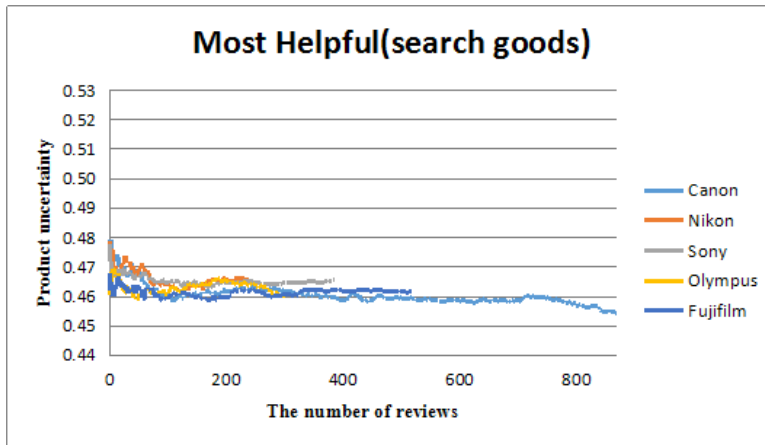


Figure 4: The Product Uncertainty of Search Goods Uncertainty Sorted by Most Helpful

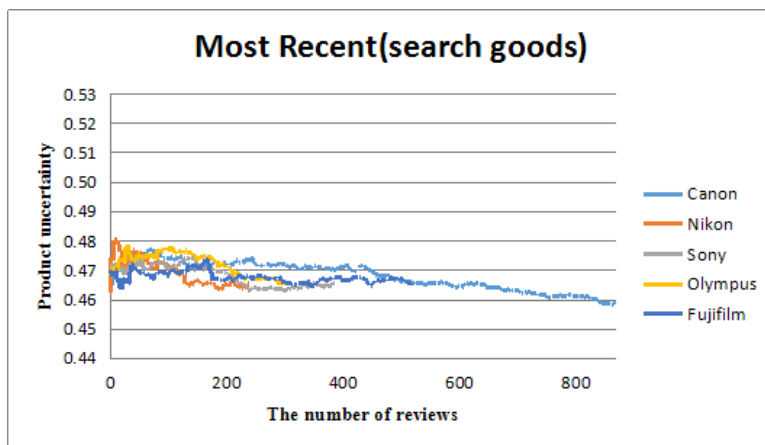


Figure 5: The Product Uncertainty of Search Goods Uncertainty Sorted by Most Recent

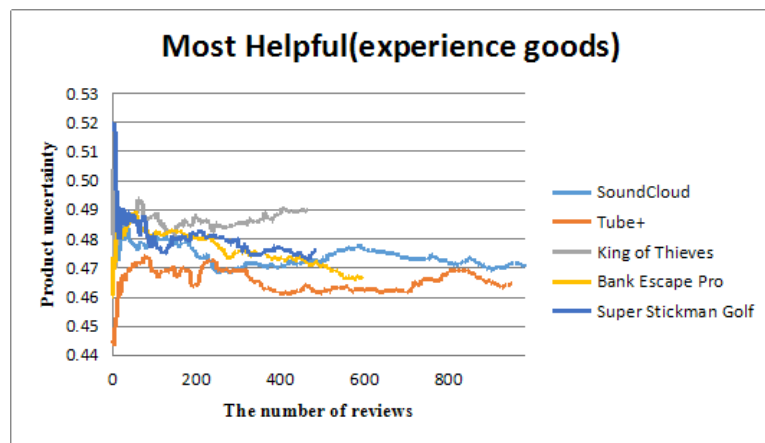


Figure 6: The Product Uncertainty of Experience Goods Uncertainty Sorted by Most Helpful

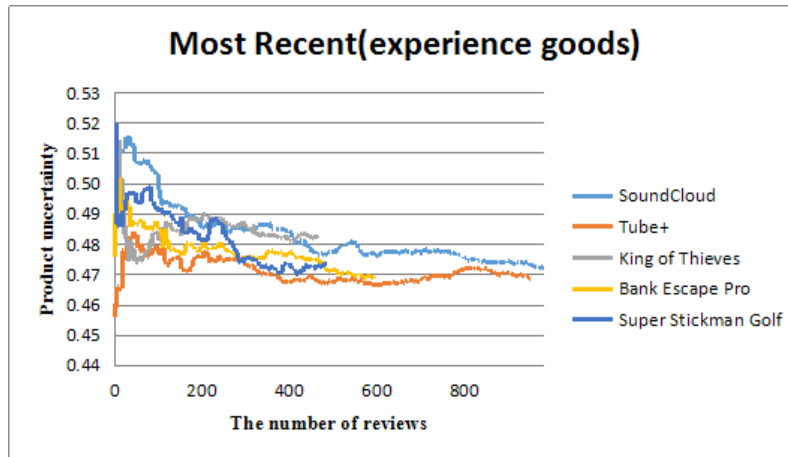


Figure 7: The Product Uncertainty of Experience Goods Uncertainty Sorted by Most Recent

In Figures 4-7, we can observe that, regardless of whether the reviews are ordered by most helpful or most recent, product uncertainty stabilizes as more reviews are considered. Additionally, when reviews are ordered by most helpful, product uncertainty steadies more quickly than that when ordered by most recent. Another interesting result obtained from these figures is that not all reviews can help consumers reduce uncertainty, particularly when all reviews are ordered by most recent. From the perspective of product type, the figures indicate that product uncertainty is relatively lower for search goods than that for experience goods. However, it is obvious that, when all reviews are ordered by helpful votes, consumers continue to need to read nearly 300 reviews to obtain stable product uncertainty; in addition, to an extent, reading reviews remains a huge and tedious job.

4. Models

In Q1, we intend to explore the impact of varied uncertainty using one review’s helpful votes.

$$helpfulvote = \theta_1 \times totalvote + \theta_2 \times rating + \theta_3 \times rating^2 + \theta_4 \times varieduncertainty + \varepsilon$$

The model derived from Mudambi & Schuff [2010] is partially modified for our purpose by adding the factor varied uncertainty from Eq. (6). In contrast to previous studies, word count is beyond our model for two reasons. First, to some extent, word count is a factor that enables consumers to decide whether to read the review or not, instead of voting for the review, while the explained variable is a helpful vote. Thus, a consumer vote has a real content value. Second, word count usually plays a role in reflecting the informativeness of the review; however, we currently have a more precise variable, varied uncertainty. The descriptive statistics of search goods and experience goods are shown in Table 2.

Table 2: Descriptive Statistics of Search Goods and Experience Goods

Variable	Search Goods		Experience Goods	
	N	Mean(SD)	N	Mean (SD)
Helpful Vote	1428	5.53(21.830)	933	4.59 (10.870)
Varieduncertainty	1428	0.00069165(0.003435269)	933	0.00211682(0.010127379)
Totalvote	1428	7.4(24.082)	933	5.84(12.398)
Rating	1428	4.16(1.223)	933	3.88(1.558)

The reviews without votes are removed; only 1428 reviews of search goods and 933 reviews of experience goods are reserved (2309 reviews of search goods and 3475 reviews of experience goods, in total). In other words, the proportions of search goods and experience goods are 61.8% and 26.8%, respectively. Such notable distinction inspires us to conduct further research with different product types as in previous studies. In addition, Table 2 shows that search goods receive more votes (7.4) and have a lower helpful votes’ proportion (75%=5.53/7.4) than experience goods (5.84 and 0.78), indicating consumers’ preference for more reviews when purchasing search goods and a tendency to vote for them more carefully. However, for experience goods, consumers appear more casual; they read fewer reviews but vote for more reviews. The ratings of search goods are higher than that of experience goods, while the standard deviation of ratings performs adversely.

Concerning Q2, motivated by the paper by Hu et al. [2014] who examined the impact of reviews on product rank with empirical data, we build on their model and consider product uncertainty as follows:

$$salerank_t = \hat{\partial}_1 \times price_{t-1} + \hat{\partial}_2 \times rating_{t-1} + \hat{\partial}_3 \times avevote_{t-1} + \hat{\partial}_4 \times uncertainty_{t-1} + \varepsilon$$

We perform the regression by lagging the one-day price and the product uncertainty. Archak et al. [2011] believe that the popularity of a product is an important factor in influencing product sales; therefore, they introduce Google trends as a control variable to indicate the popularity of products. We are inspired by Hu et al. [2014] to use average votes as an indicator instead of Google trends because Google trends reflect the mass attitude towards the product rather than the attitudes of those who are likely to purchase it. The descriptive statistics of the daily dataset are provided in Table 3.

Table 3: Descriptive Statistics of Daily Dataset

Variable	Search Goods		Experience Goods	
	N	Mean (SD)	N	Mean (SD)
Salerank	260	495.77 (714.786)	260	1244.47 (1448.166)
Price	260	227.6047 (31.69585)	260	0.800 (1.1684)
Rating	260	4.1045 (0.27458)	260	4.21153846 (0.376354104)
Avevote	260	4.85082231 (1.550428486)	260	1.86891660 (1.542767635)

As noted previously, the price of search goods is much higher than that of experience goods from Table 3, which demonstrates that price may also be a very important factor to influence consumers’ behaviors in both making purchasing decisions and voting for reviews.

5. Results

On Amazon.com’s product review website, two approaches are provided to sort reviews, namely, most helpful, ordered by the number of helpful votes, and most recent, ordered according to the reviews’ posted time. Regardless of the possibility that few consumers read reviews randomly, we assume most consumers read reviews on websites by one of the means noted above. Since the varied uncertainty by a review largely relies on its position, we conduct experiments under the following four conditions: search goods with all reviews sorted by most helpful, search goods with all reviews sorted by most recent, experience goods with all reviews sorted by most helpful, and experience goods with all reviews sorted by most recent. The conditions’ regression results are shown in Table 4. From the results, we determine that, regardless of the condition, total votes show a very high coefficient to helpful votes. We believe that total votes are a criterion for consumers to filter whether a review deserves to be read or not, while helpful votes reflect consumers’ attitude towards the review after reading it.

Table 4: Regression Output for Q1

Variable	Search Goods		Experience Goods	
	Most Helpful	Most Recent	Most Helpful	Most Recent
Varieduncertainty	0.042***	0.012*	0.024***	0.007
Totalvotes	0.959***	0.993***	0.974***	0.998 ***
Rating	0.210***	0.187***	0.033	0.021
Rating ²	-0.134***	-0.128***	0.007	0.015
Adjusted R ²	0.961	0.977	0.963	0.976

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Note that the R square values of all four regressions in Table 4 are more than 0.9, and the equations pass the F test in all four cases, which means that the helpful votes could be well explained by the following variables: interpretation rating, varied product uncertainty, and total votes.

The changes in review product uncertainty are significant and positive except for experience goods with time sequence reviews. This finding is accordance with the assumption that consumers prefer to vote for reviews that help them obtain more product attributions but that do not reduce product uncertainty. There are three reasons for the non-significant effect on the varied uncertainty of the review. First, the default review displayed on Amazon.com is ordered by voting. Second, the varied uncertainty of the review primarily depends on its position. When certain reviews with high helpful votes are ordered after the most recent, the varied uncertainty would be relatively low because most information has been contained in the earlier reviews. Third, it is more sensible to read the reviews to address search goods.

Since consumers would not read all reviews, it is very important to find valuable reviews for them. Most consumers tend to read reviews with high helpful votes to gain effective information. Consumers would stop reading when they believe they already obtain sufficient information. The analysis also provides evidence that it is easier for reviews to obtain helpful votes from experience goods than that from search goods. Above all, varied uncertainty by a review has significant impact on the helpful votes of the review, and the impact is positive.

Furthermore, for search goods, the rating and rating2 present significant effect on helpful votes, and the relationship shapes “inverted-U” curve. However, it does not suit experience goods. The result could be explained by the effect of extreme reviews that consumers focus more on extreme reviews when purchasing search goods.

Table 5 describes the output of the relationship between price, rating, average votes, product uncertainty and product rank when product uncertainty is computed by the top 10 helpful reviews, the top 20 helpful reviews, the recent top 10 reviews, the recent top 20 reviews, the total reviews, and the total reviews only considering the top 10 product attributes, respectively. Under each condition, we check the search and experience goods’ data separately.

Table 5: Model Estimates at Different Conditions for Both Types

Variable	Helpful 10		Helpful 20	
	Search Goods	Experience Goods	Search Goods	Experience Goods
Price	0.574***	0.736***	0.551***	0.746***
Rating	0.343***	0.104***	-0.286*	0.151***
Avevotes	-0.833***	-0.552**	-1.849***	-0.141**
Uncertainty	-0.023	0.578***	1.347***	0.234***
R ²	0.778	0.770	0.798	0.770

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Variable	Recent 10		Recent 20	
	Search Goods	Experience Goods	Search Goods	Experience Goods
Price	0.584***	0.656***	0.572***	0.728***
Rating	0.361***	0.271***	0.375***	0.197***
Avevotes	-0.867***	-0.17***	-0.868***	-0.07
Uncertainty	-0.065	0.297***	-0.045	0.142**
R ²	0.781	0.778	0.779	0.747

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Variable	Total		Total (Top 10 Words)	
	Search Goods	Experience Goods	Search Goods	Experience Goods
Price	0.45***	0.936***	0.632***	0.513***
Rating	0.344***	0.005*	0.213***	-0.162*
Avevotes	-0.789***	-0.076	-0.703***	-0.029
Uncertainty	0.11	0.236***	-0.188	0.494**
R ²	0.780	0.772	0.781	0.751

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

R square has been as high as 0.7 at different conditions for both product types. From the above table, we could conclude that all product uncertainties that pass the significance test show positive relationships with product rank; however, when product sales are higher, their rank would also be higher. Therefore, we could conclude that consumers tend to vote for reviews with high varied uncertainty, while they continue to prefer to purchase products with less product uncertainty from reviews. Consumers’ two main behaviors regarding e-commerce show obvious differences including a contradiction in product reviews because of the following. People are curious, particularly when purchasing products. If they want to obtain the lowest product uncertainty, the consumers only need to read product descriptions that encourage them to purchase the product. However, consumers actually continue to spend extra time and energy reading reviews from other buyers. What consumers want from reviews is reassurance, and they want to find more evidence to help them make a purchasing decision. Additionally, consumers are also afraid that important product information will be missed. Therefore, consumers want to vote for reviews that provide them new information and that allow them to purchase products with less uncertainty; this demonstrates that their voting and purchasing are self-contradictory.

It is shown in Table 5 that the uncertainty computed by the helpful top 20 reviews of search goods shows a significant effect on product rank, indicating that consumers would not read all reviews but will read the helpful 20 reviews to aid in their decision. For experience goods, uncertainty computed under all conditions show a significant effect on product rank. This finding could be explained in two ways. On the one hand, the price of search goods is much higher than that of experience goods, and it plays an intermediate role. In general, consumers would read more high-quality reviews of a product with a high price than a product with a low price. On the other hand, different types have their own characteristics. For search goods, consumers could obtain sufficient information before purchasing by reading high quality reviews. As reviews of experience goods are more subjective, consumers continue to be unable to obtain sufficient information, including after reading all the reviews, and consumers require samples before making decisions regarding experience goods. In conclusion, product uncertainty from the top 20 most helpful reviews could affect the product rank for search goods, while product uncertainty computed by every review condition could affect the product rank for experience goods.

6. Discussions

In this section, we adopt the following equation to examine the exaggerated effect of the product type and the price’s accelerating effect on the meaningful votes.

$$helpfulvote = \theta_1 \times totalvote + \theta_2 \times rating + \theta_3 \times rating^2 + \theta_4 \times varieduncertainty + \theta_5 \times price + \theta_6 \times producttype + \varepsilon$$

Table 6: Regression Outputs for Total Goods Sorted by Most Helpful Votes

Variables	Total Goods (with Price)	Total Goods (without Price)
Varieduncertainty	0.014**	0.015**
Totalvote	0.980***	0.977***
Rating	0.122***	0.134***
Rating ²	-0.055*	-0.070**
Price	-0.040***	
Producttype	0.008	-0.020***
Adjusted R ²	0.959	0.979

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 7: Regression Outputs Sorted by Most Helpful Votes for Search Goods and Experience Goods

Variables	Search Goods	Experience Goods
Varieduncertainty	0.037***	0.026***
Totalvote	0.966***	0.970***
Rating	0.228***	0.007
Rating ²	-0.145***	0.043
Price	-0.033***	-0.048***
Producttype	0.962	0.967

Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

In Table 6, when considering price, product type has little effect on helpful votes, while price shows significant impact. Moreover, only the variable price passes the T-test when simultaneously considering product type and product price. Therefore, we conclude that consumers’ behaviors on websites that include helpful votes are influenced by product price. Intuitively, one would think that when the product price is low, consumers would spend less time and be casual when they make the purchase decision. However, consumers would be more careful and cautious when the product price is too high, irrespective of the product type. We can find similar results in Table 7. In particular, even among the same product type, the product price also exerts important influence on helpful votes. To summarize, we infer that product type has no direct impact on helpful votes; however, product price plays a role as a hidden variable.

7. Managerial implication and conclusion

To our knowledge, this study first proposes that consumer online shopping behaviors are affected by continuously reading reviews, which is useful in updating their perception of a product and reducing their product uncertainty. In the theoretical analysis, we apply the Shannon entropy to address the product uncertainty; this helps us to trace the relationship between reviews. In the case study, we use real data from Amazon.com to obtain several exciting and notable findings.

The number of reviews for a product from Amazon.com can easily surpass several hundred, but not all of them receive votes. Consequently, a helpful vote recommendation system may not work well. Therefore, it is critical to

improve the efficiency of the vote recommendation system, help consumers to eliminate numerous useless reviews and effectively find valuable ones.

Previous studies address the issue by selecting useful reviews with several factors or building frameworks; however, nearly all only consider the characteristics of reviews but totally ignore the relationship between them. According to our online shopping experience, we judge a review to be helpful or not based on the how much we know about the product, which includes factors not only from the review but also from other reviews that we previously read. This process of updating our product perception by constantly reading is continuous. To understand the updating process, we apply the Shannon entropy to trace the behaviors for measuring product uncertainty. Our results confirm the varied uncertainty by a review that contributes well to helpful votes.

Several important practical implications are obtained from this paper. First, it is impossible for consumers to read all reviews before purchasing, and a recommendation system would not work effectively when the number of reviews' votes is not sufficient. According to our study, if product uncertainty, which is the factor that reflects the relationship between reviews, is considered by the reviews' recommendation system, the results could be more precise and objective. Second, the review content in the e-commerce platform could be illustrated by our technical approach to compute product uncertainty from reviews at the attributes level. This illustration can help manufacturers to quickly obtain the overall product information. If a product attribute with high uncertainty is the propaganda or headline attribute, manufacturers can offer a more convincing description on websites or find another product attribute as the promotional point, based on the performance in consumer reviews.

In addition, for search goods, the product uncertainty, which is computed by top 20 helpful vote reviews, affects their sales significantly. For experience goods, the product uncertainty in all conditions also has an important impact on product sales. Several practical meanings could be obtained from the aforementioned conclusions. First, product uncertainty serves as a filterable condition for comparing products of the same type to make better recommendations. Second, we believe consumers may prefer to read few or no reviews about apps because they do not make purchasing decisions based on the reviews' suggestion. In other words, consumers are attracted by external information. This finding implies that app sellers should launch a marketing strategy on another channel rather than spend advertising funds on maintaining their high-quality reviews.

Moreover, by applying NLTK and other text mining apparatus, manpower is dramatically liberated. Our paper is a beneficial attempt to cope with the complexity of content. This text mining process and technique could also be a reference for future studies.

We acknowledge that there are some limitations in our study. First, constrained by the technique of text mining, certain emotional words could not be judged accurately; this may affect our experiment results. In addition, implicit information in textual reviews is hardly extracted by text mining tools, and the consistency remains a thorny problem. Second, the correlation between price and helpful votes is acquired accidentally by camera and App data. Future research could obtain more categorical data to test the results. Third, we apply the number of positive words and negative words to measure product uncertainty, while the polarization of emotion words may also affect consumer perception. Finally, due to the lack of consumer preference data, we treat every consumer the same. However, such a perfect hypothesis does not actually appear. Therefore, when platforms apply product uncertainty as a factor to make recommendations for a specific consumer, they may infer his/her preference from historic behavior data and then adjust the weights of product attributes.

In summary, our paper is just a tip of the iceberg in mining the treasure of reviews with the concept of continuous reading. More interesting results could be found if future research would associate this concept with the theories of behavioral psychology.

Acknowledgements

This research is partially supported by research grants from the National Science Foundation of China under Grant Nos.71471128 and 71201115, and the Key Program of National Natural Science foundation of China under Grant No 71631003.

REFERENCES

- Archak N., A. Ghose, and P.G. Ipeirotis, "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science*, Vol. 57, No. 8:1485-1509, 2011.
- Bafna K., and D. Toshniwal, "Feature Based Summarization of Customers' Reviews of Online Products," *Procedia Computer Science*, 22:142-151, 2013.
- Chen C.C., and Y.D. Tseng, "Quality Evaluation of Product Reviews Using an Information Quality Framework," *Decision Support Systems*, Vol. 50, No. 4:755-768, 2011.

- Chong Y.L., “Mobile Commerce Usage Activities: The Roles of Demographic and Motivation Variables,” *Technological Forecasting and Social Change*, Vol. 80, No. 7:1350-1359, 2013.
- Chong Y.L., F. Chan, and K.B.Ooi, “Predicting Consumer Decisions to Adopt Mobile Commerce: Cross Country Empirical Examination between China and Malaysia,” *Decision Support Systems*, Vol. 53, No. 1:34-43, 2012.
- Clarke I., “Emerging Value Propositions for M-commerce,” *Journal of Business Strategies*, Vol. 18, No. 2:133-148, 2001.
- Dimoka A., Y. Hong, and P.A. Pavlou, “On Product Uncertainty in Online Markets: Theory and Evidence,” *MIS Quarterly*, Vol. 36, No. 2:395-426, 2012.
- Duan W., B. Gu, and A.B. Whinston, “Do Online Reviews Matter?—An Empirical Investigation of Panel Data,” *Decision Support Systems*, Vol. 45, No. 4:1007-1016, 2008.
- Forman C., A. Ghose, and B. Wiesenfeld, “Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets,” *Information Systems Research*, Vol. 19, No. 3:291-313, 2008.
- Ganu G., Y. Kakodkar, and A. Marian, “Improving the Quality of Predictions Using Textual Information in Online User Reviews,” *Information Systems*, Vol. 38, No. 1:1-15, 2013.
- Hong Y., and P.A. Pavlou, “Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents,” *Information Systems Research*, Vol. 25, No. 2:328-344, 2014.
- Hu H.W., Y.L. Chen, and P.T. Hsu, “A Novel Approach to Rate and Summarize Online Reviews according to User-Specified Aspects,” *Journal of Electronic Commerce Research*, Vol. 17, No. 2:132-152, 2016.
- Hu N., N.S. Koh, and S.K. Reddy, “Ratings Lead You to the Product, Reviews Help You Clinch It? The Mediating Role of Online Review Sentiments on Product Sales,” *Decision Support Systems*, Vol. 57:42-53, 2014.
- Huang L., X. Lu, and S. Ba, “An Empirical Study of the Cross-channel Effects between Web and Mobile Shopping Channels,” *Information & Management*, Vol. 53, No. 2:265-278, 2016.
- Kiang M.Y., T.S. Raghu, and K.H. Shang, “Marketing on the Internet—Who can Benefit from an Online Marketing Approach?” *Decision Support Systems*, Vol. 27, No. 4:383-393, 2000.
- Kiang M.Y., and R.T.H. Chi, “A Framework for analyzing the potential benefits of Internet marketing,” *Journal of Electronic Commerce Research*, Vol. 2, No. 4:157-163, 2001.
- Katz E., and F. Lazarsfeld, “Personal Influence: The Part Played by People in the Flow of Mass Communications,” Free Press: New York, 1955.
- Kim S.M., P. Pantel, T. Chklovski, and M. Pennacchiotti, “Automatically Assessing Review Helpfulness,” *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pp. 423-430, Sydney, July 2006.
- Knight F.H., “Risk, Uncertainty and Profit,” Mass: Houghton Mifflin, 1921.
- Kolomiiets A., N. Dens, and P. De Pelsmacker, “The Wrap Effect in Online Review Sets Revisited: How Perceived Usefulness Mediates the Effect on Intention Formation,” *Journal of Electronic Commerce Research*, Vol. 17, No. 4:280-288, 2016.
- Li X., and L.M. Hitt, “Self-selection and Information Role of Online Product Reviews,” *Information Systems Research*, Vol. 19, No. 4:456-474, 2008.
- Liu Y., X. Huang, A. An, and X. Yu, “Modeling and Predicting the Helpfulness of Online Reviews,” *Proceedings of the 8th IEEE International Conference on Data Mining*, pp. 443-452, Pisa, December 2008.
- Martin L., and P. Pu, “Prediction of Helpful Reviews Using Emotions Extraction,” *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pp. 1551-1557, Québec, July 2014.
- Martín S., B. López-Catalán, and M. Ramón-Jerónimo, “Factors Determining Firms’ Perceived Performance of Mobile Commerce,” *Industrial Management & Data Systems*, Vol. 112, No. 6:946-963, 2012.
- Montemurro M.A., and D.H. Zanette, “Entropic Analysis of the Role of Words in Literary Texts,” *Advances in Complex Systems*, Vol. 5, No. 1:7-17, 2002.
- Mudambi S.M., and D. Schuff, “What Makes a Helpful Review? A Study of Customer Reviews on Amazon.com,” *MIS Quarterly*, Vol. 34, No. 1:185-200, 2010.
- Netzer O., R. Feldman, J. Goldenberg, and M. Fresko, “Mine Your Own Business: Market-structure Surveillance through Text Mining,” *Marketing Science*, Vol. 31, No. 3:521-543, 2012.
- Putthividhya D.P., and J. Hu, “Bootstrapped Named Entity Recognition for Product Attribute Extraction,” *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1557-1567, Edinburgh, July 2011.
- Quan C., and F. Ren, “Unsupervised Product Feature Extraction for Feature-oriented Opinion Determination,” *Information Sciences*, Vol. 272:16-28, 2014.

- Shankar V., and S. Balasubramanian, "Mobile Marketing: A Synthesis and Prognosis," *Journal of Interactive Marketing*, Vol. 23, No. 2:118-129, 2009.
- Shannon C.E., "A Note on the Concept of Entropy," *Bell System and Technical Journal*, 27:379-423, 1948.
- Wang J.H., E.C. Malthouse, and L. Krishnamurthi, "On the Go: How Mobile Shopping Affects Customer Purchase Behavior," *Journal of Retailing*, Vol. 91, No. 2:217-234, 2015.
- Wong C.H., H.S. Lee, Y.H. Lim, B.H. Chua, and G. Tan, "Predicting the Consumers' Intention to Adopt Mobile Shopping: An Emerging Market Perspective," *International Journal of Network and Mobile Technologies*, Vol. 3, No. 3:24-39, 2012.
- Xu K., S.S. Liao, J. Li, and Y. Song, "Mining Comparative Opinions from Customer Reviews for Competitive Intelligence," *Decision Support Systems*, Vol. 50, No. 4:743-754, 2011.
- Yang K., "Determinants of US Consumer Mobile Shopping Services Adoption: Implications for Designing Mobile Shopping Services," *Journal of Consumer Marketing*, Vol. 27, No. 3:262-270, 2010.
- Yang K., and H.Y. Kim, "Mobile Shopping Motivation: An Application of Multiple Discriminant Analysis," *International Journal of Retail & Distribution Management*, Vol. 40, No. 10:778-789, 2012.
- Ye Q., R. Law, and B. Gu, "The Impact of Online User Reviews on Hotel Room Sales," *International Journal of Hospitality Management*, Vol. 28, No. 1:180-182, 2009.
- Yu J., Z.J. Zha, M. Wang, and T.S. Chua, "Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews," *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 1496-1505, Portland, June 2011.
- Zhang K., S.J. Zhao, C. Cheung, and M. Lee, "Examining the Influence of Online Reviews on Consumers' Decision-making: A Heuristic-systematic Model," *Decision Support Systems*, Vol. 67:78-89, 2014.
- Zhang R., and T.T. Tran, "Helping E-commerce Consumers Make Good Purchase Decisions: A User Reviews-based Approach," in Babin G., Kropf P., Weiss M. (eds) *E-technologies: Innovation in an open world*, Springer, Berlin, pp. 1-11, 2009.