2012

Online reviews and consumers' willingness to pay: the role of uncertainty

Yinglu Wu
Louisiana State University and Agricultural and Mechanical College

ONLINE REVIEWS AND CONSUMERS' WILLINGNESS TO PAY: THE ROLE OF UNCERTAINTY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Interdepartmental Program
in
Business Administration (Marketing)

by
Yinglu Wu
B.S., Hubei University, 2005
M.S., Louisiana State University, 2012
December 2012

To my mom, who always loves me more than I do

ACKNOWLEDGEMENTS

To Dr. Jianan Wu, my chair and mentor, thank you so much for everything you have done. Words cannot express how grateful I am for having been able to work under your guidance. When I joined the PhD program, I was not quite sure about my direction; it was you who showed me the way to my passion for marketing and research. Not only have you taught me the rigor of research, but you have also demonstrated to me a lifelong attitude of professionalism and boundless curiosity. I'm also grateful for your extreme patience. Despite the many (I mean MANY) mistakes I made along the way, you did not give up on me. I could not have had a better mentor and you will always be a role model for me.

To Dr. Bill Black, Dr. Alvin Burns, and Dr. Tim Chandler, thank you for serving on my committee and for all the advice and encouragement you provided to help me complete my dissertation. To the marketing, management, and statistics professors who taught me over the past six years, thank you for imparting the knowledge that shapes my view of research.

To all my friends at LSU, I would not have been able to complete the PhD program without your support. To the Garritys-Carolyn, Dan, Maddie, and Keenan-you are my family in the US; to Mazen, Nobu, Kate, Dora, Eric, Stephanie, Jacob, Jie, Linda, Mousumi, Anna, and Yanna, I'm so blessed to have met every one of you. The time we had together at LSU is such a wonderful part of my life.

Lastly, to my dear mom, Andi Min, thank you for your sacrifice and dedication. I couldn't have done anything in my life without your love and support!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Empirical studies of online reviews have found that valence (average rating) has a consistently positive impact on consumers' willingness to pay (WTP), but volume does not. Although two studies tried to explain this phenomenon using different perspectives (Wu and Ayala, 2012; Sun, 2012), neither study can fully accommodate the consumer behaviors observed by the other. This dissertation adopts a theoretical framework that can explain the consumer behaviors observed in both studies as well as the varying influence of review volume at the individual level. Specifically, several studies were conducted to investigate the relationship between bidirectional online seller reviews (e.g., the eBay review format) and consumers' WTP.

Essay 1 provides an extensive review of studies that investigate online consumer reviews at the market, product, firm, consumer, and message level; special attention is given to the outcomes of consumer reviews for both products and sellers. In addition, this essay establishes the importance of the current research topic.

Essay 2 combines economic and behavioral theories of decision-making under uncertainty to develop a theoretical framework. The framework proposes that review volume and valence influence a consumer's WTP through a weighting function of outcome probability. Consumers with different preferences towards uncertainty will have different preferences toward review volume, and for some consumers, such preference can change depending on the review valence. Based on this conceptualization, the framework reconciles the current literature by explaining the inconsistent influence of review volume both across and within individuals. The internal validity of the framework is tested with an experiment and analyses carried out at the individual level provide strong support for the proposed conceptual model.

Essay 3 establishes the relevance of this research for managers by applying the framework to real market data. Due to the nature of the transactional data, a finite mixture model is used to estimate the weighting function, and hypotheses are tested at the group instead of the individual level. A simulation study demonstrates the validity of using a finite mixture model to estimate the weighting function and classify groups. The results of the hypotheses testing provide adequate support for the framework.

ESSAY ONE. AN APPRAISAL OF ONLINE USER REVIEWS

INTRODUCTION

By nature, online purchases involve much more uncertainty than offline purchases. Online reviews, a digital form of consumer word-of-mouth, provide a useful tool for reducing the uncertainty of purchases. Ample evidence shows that online reviews have become an important component of consumers' purchase decisions. Nielsen's 2010 online shopping report reveals that online reviews and peer recommendations have become a key factor that influences consumers' purchases, especially those of electronics, cars, and travel. Forty percent (40%) of online consumers indicate that they will not buy electronics without reading online reviews first. In Nielsen's more recent report on advertising trustworthiness (2012), online consumer opinions is ranked as the second-most trustworthy and second-most relevant form of advertising when searching for information about products, trailing only recommendations from the consumer's personal network. Academic studies also confirm the importance of online reviews. For example, Bronner and de Hoog (2010) found that tourists rated consumer-generated review sites as more up-to-date and useful than market-generated sites (2010). Utz at al. (2012) found that consumer reviews of an online retailer are a more important indicator of trustworthiness than the overall store reputation.

In contrast to traditional word-of-mouth, online reviews can be massive in scale. The assessment of a product or seller is no longer limited to one or two customers' experiences; those assessments may come from hundreds, thousands, or even millions of customers. On the other hand, in offline word-of-mouth communication, a consumer typically knows the communicator and is able to judge the quality of the assessment based on that knowledge. Such personal knowledge about online communicators is generally missing. Because of these unique characteristics, online reviews have drawn a great deal of attention from researchers. Despite the huge efforts that are devoted to this topic, we still lack a deep understanding of the mechanisms that tie online reviews to consumer decisions, product or firm performance, and market efficiency. My research goal is to explore how online reviews influence consumer's price decisions and provide insights for managers to better utilize online reviews to increase their firms' marketing performance. To reach this goal, I conducted an extensive literature review to ascertain current knowledge about online reviews.

Method

The scope of this review is limited to consumer-generated online reviews about products, individual sellers, and firms. The purpose of my research is to study the impact of massive consumer reviews on consumer decisions, so I exclude research (1) that focuses on consumer-generated content in the form of blogs or social network platforms, because the impact of social ties is not relevant to the current research, and (2) that examines objective third-party reviews, such as reviews from consumer reports or professional organizations.

Following the call for multi-disciplinary research on e-commerce (Taylor and Strutton, 2010), I reviewed research in the following disciplines: marketing, management, information science, and economics. I selected the top 20 journals ranked by ISI impact factor and the top 20 journals ranked by ISI 5-year impact factor in the categories of business, management, information science & library science, and economics. The final list included 57 journals, each of which I reviewed from 2000 to the present. The list of journals is shown in the appendix.

In the rest of Essay One, I summarize the current literature and explain my research motivations. First, I briefly review the areas of research that involve online consumer reviews; second, I provide a more detailed review of the outcomes of consumer reviews, for both products and sellers/firms; last, I discuss the studies that motivate my research topic and the structure of my dissertation.

ONLINE USER REVIEWS

Websites commonly use two types of review systems. The first is a star rating system, by which a consumer can rate a seller or a product using a Likert scale; for example, Amazon uses a 5-star review system, with 1 being the lowest value and 5 being the highest. The second is a bidirectional review system that assumes that a consumer will provide a positive review if satisfied and a negative one if not, such as eBay's review system. Most review systems provide statistical summaries of the reviews: review volume is the number of reviews received for a specific seller or product and review valence is the average of the review ratings. Even though many systems do not directly report the variance of reviews, it can be inferred by the consumers in various ways, for example, by looking at the distribution of reviews. Examples of these two review systems are shown in Figures 1.1.

In the following review, I organize the research based on focus and topic. Studies of online reviews have very different emphases and scopes. Some studies focus on market-level outcomes, such as the characteristics of review distributions in various markets and the effectiveness of employing review systems to improve market efficiency. Some studies focus on the product/firm level, exploring the generation and consequence of reviews for a specific product or seller. Some research looks closely at the consumer or message level, studying what factors motivate a consumer to post online reviews or what types of messages persuade a consumer.

5-star Review System (Source: Amazon.com)

Bidirectional Review System (Source: eBay.com)

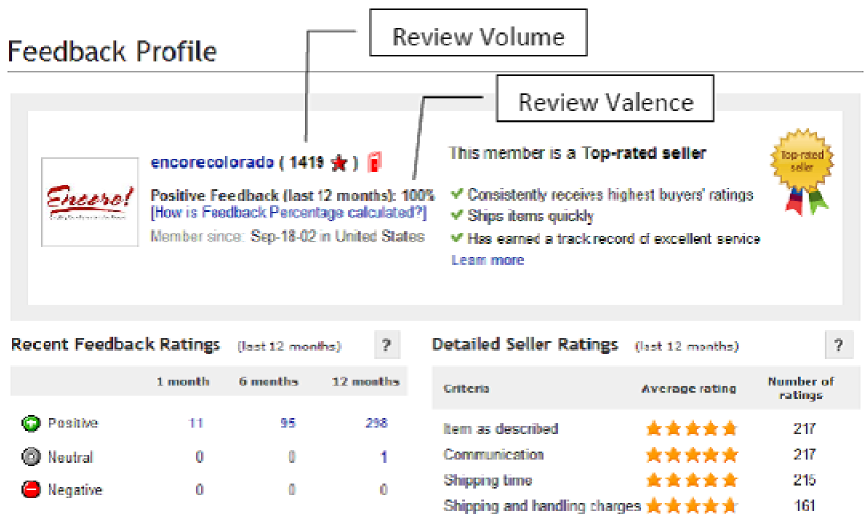


Figure 1.1 Examples of Two Review Systems

Market-Level Research

One stream of literature provides insights on the design of review systems. Through modeling, experiments, and online empirical studies, these studies identify the conditions under which review systems are useful for generating efficient economic outcomes. Bakos and Dellarocas (2011) utilize game theory to demonstrate that an online reputation system is very important for a market in which adverse selection exists; a reputation-based system helps sellers and buyers learn about each other, benefitting both participants with high quality. They also suggest that reputation systems are very important for the professional services market, where other endurance mechanisms may cost a lot and the service outcome depends more on the type of rather than the effort by the seller. Introducing review systems can improve buyers' well-being and their willingness to trade in that market (Yang et al., 2007), and the larger the impact of the review system on the transaction outcomes (rewards for positive reviews and punishments for negative reviews), the more likely the sellers will be honest (Zhou et al., 2008). Within a repeated-game setting, Yang et al. (2007) conclude that the mere existence of a review system, no matter how simple, helps improve market performance. Dellarocas (2005) also find that a simple binary review profile, such as eBay's review mechanism, can stimulate maximum market efficiency. Kumar and Benbasat (2006) argue that allowing a consumer to provide a review not only improves the consumer's perception of the website functionality, but also strengthens the social connection between the website and the consumer.

Even though review systems can enhance market-level honesty, dishonest behavior can still exist. Yang et al. (2007) demonstrate that there is a correlation between a seller's tendency to cheat and her reputation; that is, the more the seller tends to cheat, the more likely she will build a good reputation. Moreover, dishonest sellers can manipulate reviews at a relatively low cost. Since reviews are anonymous, dishonest sellers can submit good reviews for themselves, and bad sellers can still participate in the market by starting over with a new ID. Some studies have identified the conditions that enhance or limit the effectiveness of a review system in promoting seller honesty. Zhou et al. (2008) find that the effectiveness of review systems can be limited if buyers are not motivated to review sellers after transactions. Aperjis and Johari (2010) examine the number of past transactions that should be used to calculate a seller's reputation. They find that calculating the seller's reputation over a longer duration of time and a larger number of transactions is more likely to make patient sellers truthful but less likely to make high-quality sellers truthful. Finally, Bolton et al. (2008) suggest that encouraging market-level competition can increase the effectiveness of a review system by building trust, and trustworthiness, in the market.

To specifically deal with fraud in review systems, Abbasi et al. (2008) propose a stylometric method for identifying a trader by analyzing he writing style of the feedback comments she posts. You et al. (2011) also propose a set of indicators that can detect fake transactions and puppet buyers on consumer-to-consumer auction sites. By comparing regular and collusive transactions on a large Chinese auction site, they find that buyers for collusive transactions are usually more active and have a shorter history than regular buyers, collusive transaction items are on average less valuable than regular transaction items, and puppet buyers are more likely to present detailed comments for collusive sellers.

Product- and Firm-Level Research

Research at the product/firm level has focused on four areas: (1) antecedences of reviews, (2) changes of review structure overtime, (3) outcomes of reviews, and (4) marketing strategies that incorporate online reviews. I do not summarize review outcomes in this section, providing a more detailed discussion later in the essay.

Antecedences of reviews. Studies have identified factors that influence the volume and the valence of product reviews.

<u>Factors that influence review volume.</u> One stream of literature explores the factors that may influence the generation of online reviews, which has been shown to be associated with market factors such as popularity (Dellarocas et al., 2010) and sales (Moe and Trusov, 2011; Dellarocas et al., 2010; Feng and Papatla, 2011), firms' strategies such as advertising spending (Feng and Papatla, 2011), and existing reviews for the product (Moe and Trusov, 2011,

Dellarocas et al., 2010).The most intuitive factor that influences the propensity of reviews is sales, since the greater the product sales, the more experience consumers have with the product, and the more likely they are to post reviews for that product (Moe and Trusov, 2011; Feng and Papatla, 2011). Examining market-level data, Dellarocas et al. (2010) find that consumers prefer to post reviews for movies that are less popular and less successful; they also like to post reviews for movies that have already accumulated many comments. Correspondingly, the authors observe a U-shaped relationship between review posting volume and a movie's box office revenue, in which more reviews are posted for either very obscure movies or high box-revenue movies. Feng and Papatla (2011) find that the amount spent on advertising for an automobile brand is negatively associated with the number of reviews posted in the same year. Comparing two data sets collected in 2001 and 2008, respectively, Chen et al. (2011) find that, in general, there are more reviews posted for products of extremely low or extremely high quality. During the early stages of internet use, the price of a product negatively influences the aptness of reviews for that product. As internet use becomes more common among consumers, price exhibits a U-shaped relationship with review volume: more reviews are observed for products that either have extremely low or extremely high prices.

Factors that influence review valence. Li and Hitt (2010) propose that consumer reviews should reflect their evaluation of not only product quality but also product value. In a study of reviews for cameras, they find that, when controlled for camera quality, the average of review ratings will rise by 0.16 (on a 1-10 scale) if the camera price drops by 20% and 0.36 if the price drops by 40%.In a study of automobile reviews, Chen et al. (2011)find that, although price has a negative but statistically insignificant influence on review ratings, it has a U-shaped relationship with review valence in the early stages of internet usage, in which lower or higher priced products tend to have higher ratings than moderately priced products. For experiential products, review valence is found to be positively related to the number of product users (Yang and Mei, 2010).Koh et al. (2010) study the influence of culture on online review valence. In a review of ratings for movies by consumers from China, Singapore, and the U.S., they find that Chinese consumers are less likely than American consumers to provide extreme ratings. Correspondingly, they observe a U-shaped distribution of review valence on American movie review sites, but a bell-shaped distribution of review valence on Chinese movie review sites.

Review evolvement. Studies also examine review evolvement, most of them using longitudinal data to capture the progression of review profiles over time. Li and Hitt (2008) attribute the changes in product reviews over time, which usually follow a falling trend, to the fact that early reviewers, who are also initial buyers of a product, self-select the products they believe they are more likely to enjoy, and hence their evaluations tend to be more positive. The opinions of earlier buyers, however, do not necessarily reflect those of later buyers. Li and Hitt also find that when consumers use product reviews to form purchase decisions, they do not fully correct the self-selection bias. As a result, later buyers' reviews tend to be lower than early buyers', and the majority of the reviews follow a declining trend over time.

Moe and Trusov (2011) find that increases in review valence tend to incite new negative reviews and discourage the subsequent posting of extremely positive reviews; increases in variance among existing reviews discourage the posting of extreme reviews; and an increase in review volume increases the posting of reviews in general. Using book review data from Amazon and controlling for book quality, Hu and Li (2011) find that later reviews for a book tend to deviate from previous reviews. That deviation is more likely if the later reviews mention the earlier reviews, the existing reviews have a large volume or a small variance, and the book is not popular among consumers.

Firm's marketing strategy. Given the overwhelming evidence that online reviews contain valuable information for consumers as well as companies, studies propose various ways that companies can incorporate online product reviews into their marketing strategies. Chen and Xie (2008) develop a normative model to show that firms should incorporate consumer reviews when developing their communication strategies. Companies' responses to online consumer reviews should take into consideration the relative size of the expert consumer segment and the cost of the product. Companies should release more product attribute information in response to consumer reviews if the product cost is low or the expert consumer segment is large, but reduce the amount of information if the product cost is high or there are not enough expert consumers. Chen and Xie (2008) also suggest that companies should be cautious about providing consumers the option to leave reviews on their websites. Such a feature benefits products when the novice consumer segment is large, but can hurt the company when the expert consumer segment is large. Several studies also propose marketing research methods or models that retrieve information from online consumer reviews to provide insights for companies' positioning (Lee and Bradlow, 2011) and product strategies (Decker and Trusov, 2010).

Consumer- and Message-Level Research

Studies that focus on the consumer level explore individual characteristics that lead to different behaviors in terms of posting and using online reviews. Many researchers also look at individual review messages and identify qualities that make one message more persuasive than another.

Loyalty to review systems. Wang et al. (2010) find that people tend to continuously use an online review system if they have a high propensity to learn about and adopt online review systems, and if they view review systems as very relevant to their personal needs and interests. In a survey of online users, Awad and Ragowsky (2008) find that gender plays a role in the perception of the quality of a review system and of trust towards a website. Men view a review system as having better quality if it provides many opportunities for the consumer to post opinions, if there is a high volume of responses, and if others participate. For women, the response from and participation of others is very important, but the opportunity to post opinions is negatively associated the quality of the review system. For men and women, the helpfulness and ease of use of a review system positively influences their trust of the website and hence their intention to use it, but women place more weight on ease than men. Park and Lee (2009) propose that a consumer will use online reviews more and be more likely to be influenced by them if she is more susceptible to interpersonal influence and has more online shopping experience. They also find that the relationship between these personal characteristics and online review usage behavior is stronger for Korean consumers than for U.S. consumers.

Review posting behavior. Additional studies explore what types of consumers are more likely to post reviews. Usually posting behavior is associated with a consumer's personal characteristics and experience with the purchase. Many studies have documented that consumers who have the highest and lowest satisfaction levels are more likely to post reviews, which leads to an under-reporting bias (Koh et al., 2010).

Henning-Thurau et al. (2009) closely examine the underlying motives of consumers who post opinions. In an analysis of comments posted on a German opinion website, they find that concern for other customers, the social benefit of affiliating with a virtual community, a desire for positive recognition from others, the economic rewards from website operators, and a need for advice are the dominant motives. These motives are associated with the frequency of a consumer’s visits to the website and the number of comments she wrote. They also suggest that firms can segment consumers based on their motives for posting opinions online.

Review adoption. Using a simulation of an online auction site, Wolf and Muhanna (2011) find that consumers usually focus on review valence information and underweight review volume. Moreover, they find that this bias is more prevalent for the star-scale review format, such as Amazon’s, than for the binary review format, such as eBay's. Some studies suggest that different consumers will process review information differently. For example, Lee et al. (2008) find that high-involvement consumers tend to be influenced by negative reviews that have high quality; however, low-involvement consumers tend to conform to negative reviews regardless of review quality. Park and Kim (2008) propose that experts like to process information about product attributes and infer the benefit based on their knowledge, but novices like to process information that directly discloses product benefits. Hence, reviews focusing on product attributes have more impact on experts’ purchase intentions, while comments focusing on product benefits have more impact on novices’ purchase intentions.

Review message persuasiveness. As mentioned above, consumers provide reviews of products and sellers for various reasons; their backgrounds also vary widely in terms of interest and knowledge. Therefore, readers do not perceive reviews as equal in quality or credibility. Many studies show that consumers do read more than the statistical summary of the review profile; they also will read individual reviews and heed the authors. DeMaeyer and Estelami (2011) document that consumers trust experts’ opinions more for goods, but users’ testimonials more for services. Naylor et al. (2011) argue that consumers’ perceptions of the similarities between themselves and the reviewers will impact how much they are persuaded by the reviews. When information about a reviewer is missing, readers will infer that the ambiguous reviewer is similar to them; hence, consumers tend to agree with reviews posted by ambiguous reviewers more than with reviews posted by dissimilar reviewers. Lee et al. (2008) find that the influence of negative product reviews on consumers’ attitudes towards a product is moderated by the quality of the review, as measured by relevancy, reliability, understandability, and sufficiency. Kim and Gupta (2011) study the emotional expression in review messages, and find that consumers tend to attribute negative emotions to a reviewer's irrational dispositions; therefore, the expression of negative emotions in a negative review decreases its persuasiveness. However, the expression of positive emotion in a positive review does not improve the consumer's evaluation of the target.

Consumers do not just care about the review content for product information; they also care about the content of reviews for online sellers. Pavlou and Dimoka (2006) conducted a large-scale content analysis of reviews posted on eBay, finding that the review text generated significant economic value beyond the numerical ratings. After controlling for a seller's numerical ratings, they find that reviews that comment on a seller's outstanding/abysmal benevolence and outstanding/abysmal credibility will influence consumers' trust of the seller and, as a result, impact the price premiums charged by the seller.

Some studies suggest that consumers may choose to trust and rely on only parts of a review. Yang and Mei (2010) find that for experiential products such as video games, consumers tend to trust comments about search attributes but not high-level experiential attributes. Finch (2007) finds that on eBay, reviews about the quality of a seller's services such as delivery, communication, and problem solving are very important for low-risk products, or new products of low value. However, reviews about the quality of the product, such as its condition, and whether the product is exactly as described by the seller are very important when there high risks associated with the product, for example, used products or high-priced products.

Websites like Amazon.com and Epinions.com also provide rating systems for the review itself. Amazon.com lets consumers indicate whether they feel a review is helpful or not, and consumers can also comment on reviews provided by others. Epininons.com, a website that allows consumers to review various products, uses two ratings to help consumers identify high-quality reviews. The first rating assesses the content of the review: a consumer can rate each review as not helpful, somewhat helpful, helpful, or very helpful. The second rating assesses the source of the review. Each reviewer has a profile that lists all of the reviews she has provided, and a consumer can choose to “trust” the reviewer or “block” the reviewer. Studies also analyze the content of reviews to determine what types of reviews score highest on the helpfulness rating. Message content, such as one-sided vs. two-sided argument and evidence presentation, and written style, such as readability, comprehensiveness, and language intensity, are found to be associated with the helpfulness ratings of reviews (Korfiatis et al., 2011; Li and Zhan, 2012). Mudambi and Schuff (2010) also propose that while the extremity of a review and the depth of a review influence the helpfulness rating, these relationships are moderated by whether the target product is a search or experience product.

ONLINE USER REVIEWS AND THE OUTCOMES

In this section, I discuss the outcomes of online consumer reviews in detail. Reviews can be written about products or sellers/firms. Research shows that product reviews usually provide information about product attributes, functionality, and benefits (Park and Kim, 2008); seller reviews usually disclose information about product quality, such as product conditions, as well as seller quality, such as delivery and communication (Lei, 2011). The studies discussed in this section focus on product- or firm-level review characteristics and outcomes. Specifically, many studies investigate statistical summaries of reviews, such as review volume, valence, and variance. Most of the studies explore the influence of online reviews on aggregate consumer behavioral outcomes, for example, product sales (Gilkeson and Reynolds, 2003; Chevalier and Mayzlin, 2006; Li and Hitt, 2008; Chen et al., 2011), sales price (Melnik and Alm, 2002; Zhang, 2006; Reiley et al., 2007; Wu and Ayala, 2012), product revenue (Basuroy et al., 2003; Liu, 2006; Duan et al., 2008; Moon et al., 2010), and firm's financial performance (Chen et al., 2012; Tirunillai and Tellis, 2012). A few studies investigate the influence of online reviews on consumer attitudinal outcomes such as preference (Lee and Lee, 2009; Khare et al., 2011) and trust (Ba and Pavlou, 2002). I summarize the studies on product reviews and on seller reviews separately.

Reviews for Products

There are twenty articles that directly test the consequences of product reviews. Seven of those studies focus on the motion picture industry and examine movie sales and revenues. Other studies focus on software, books, video games, digital cameras, beauty products, etc. The most extensive study is one by Tirunillai and Tellis (2012), which involves 15 firms across 6 markets. Table 1.1 shows that, in terms of sales and revenue, review valence has more consistent influence than volume and variance; however, in terms of companies' financial performances, such as stock market return, volume seems to have more influence than valence. While Tirunillai and Tellis (2012) find that review valence has no impact, Chen et al.(2012) find that it is changes in the review valence, not the absolute valence, that affects firms' stock returns.

Some authors also suggest the importance of looking at interactions between review statistics and other possible moderators. Sun (2012) find that, for online book sales, there view valence interacts with review variance, so that when valence is low, higher variance leads to higher sales. Khare et al. (2011) demonstrate the possibility of interactions among review valence, variance, and volume in forming consumer preferences. While many studies find that negative reviews have more impact than positive reviews on sales and revenue (Basuroy et al., 2003; Chen et al., 2011), Clemons et al. (2006) find that for beer, a frequently purchased product, high-end reviews actually carry more weight than low-end reviews. Increasing the variety of the products in one category can also weaken the relationship between product reviews and sales (Zhou and Duan, 2011).Another important aspect of product that needs to be considered is popularity. Duan et al. (2009) find that review valence does not influence the adoption of popular software, but it has a significant impact on adoptions of less popular (niche) software. Similarly, Zhu and Zhang (2010) find that review valence and variance only impact sales of less popular video games. Park et al. (2011) suggest looking beyond product reviews in a single market, because consumers can visit different websites to obtain review information for the same product. The authors find that the relationship between review valence and sales on one website is influenced not only by the volume of reviews accumulated on that website, but also by the volume of reviews for the same product on other websites. The detail results of these twenty studies are listed in Table 1.2.

Table 1.1 The Impact of Review Volume, Valence, and Variance

Dependent Variable	Article	Review Volume	Review Valence*	Review Variance
Sales	Chevalier and Mayzlin, 2006		+	
	Clemons et al., 2006		+	+
	Li and Hitt, 2008	+	+	
	Duan et al., 2009		+ or NS	
	Chintagunta et al., 2010	NS	+	NS
	Moon et al., 2010		+	
	Zhu and Zhang, 2010	+	+ or NS	− or NS
	Chen et al., 2011	+	NS	
	Park et al., 2011		+	
	Sun, 2012	+	+	
Revenue	Basuroy et al., 2003	+, −, or NS		
	Liu, 2006	+	NS	
	Duan et al., 2008	+		
	Moe and Trusov, 2011	NS	+	NS
Financial Performance	Chen et al., 2012		NS	
	Tirunillai and Tellis, 2012	+	NS	

* For studies that also report the valence of negative reviews, I only summarize the impact of positive review valence here. Table 1.2 provides the full results of review valence.

Table 1.2 Summary of Product Review Outcomes

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Basuroy, Chatterjee, and Ravid, 2003	Number of positive (negative) reviews, percentage of positive (negative) reviews, and review volume	Revenue	Movie	Variety and Baseline.Hollywood.com	Review valence (both positive percentage and negative percentage) influence revenue. Negative review number influences revenue more than positive review number, but the influence of negative review number diminishes over time. Review volume has mixed influence on revenue. In different weeks after the movie is released, it has either positive, negative or no impact on revenue.
Chevalier and Mayzlin, 2006	Review valence (5-star scale)	Sales	Book	Amazon Barnesandnoble.com	Increase in review valence leads to increase in relative sales. The impact of negative (1-star) reviews is larger than positive (5-star) reviews.
Clemons, Gao, and Hitt, 2006	Review valence and variance	Sales	Beer	Ratebeer.com	Both review valence and variance are positively related to future sales. High-end ratings are weighted more than low-end ratings, because beer is a repeat purchase product.
Liu, 2006	Review volume and percentages of positive (negative) messages	Revenue	Movie	Yahoo! Movies	WOM activities are most active during a movie's prerelease and opening week. WOM offers significant explanatory power for both aggregate and weekly box office revenue, especially in the early weeks after a movie opens. Most of this power comes from the volume of WOM (through awareness), not from its valence (through attitude).

Table 1.2 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Duan, Gu, and Winston, 2008	Review valence, volume, and revenue	Review volume and revenue	Movie	Variety, Yahoo! Movies, and Box Office Mojo	Separate the effect of online WOM as a precursor and an outcome of retail sales. Both a movie's box office revenue and WOM valence significantly influence WOM volume; volume in turn leads to higher box office revenue
Li and Hitt, 2008	Review volume and valence (5-star scale)	Sales	Book	Amazon	Sales are positively related to review volume and valence.
Duan, Gu, and Whinston, 2009	Review valence (5-star scale)	Product adoption	Software	CNET Download.com	Product ratings have no impact on users' choice of popular software, and have a significant positive impact on the adoption of less popular products.
Lee and Lee, 2009	Review valence and variance	Purchase intention and preference for product	Windows Vista and movie	Experimental survey	Review valence and variance moderate the impact of product attributes: quality and preference on consumers' purchase intentions. Review valence and variance moderate the impact of perceived quality on product preferences.
Lee and Youn, 2009	Review valence (positive vs. negative) and platform	Product recommendation	Apartment	Experiment	Review valence influence recommendation intent. When review is positive, the impact is moderated by the platform of reviews.
Chintagunta, Gopinath, and Venkataraman, 2010	Review volume, valence (13-item scale), and variance	Sales	Movie	Yahoo! Movies	Review valence has positive impact on sales. Neither volume nor variance has impact on sales.

Table 1.2 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Moon, Bergey, and Iacobucci, 2010	Review valence	Revenue	Movie	Rotten Tomatoes and Yahoo! Movies	Critics' ratings significantly influence movie revenue during the opening week while amateurs' do not. Amateurs' ratings influence movie review in the later weeks only when they are supported by heavy ad spending.
Zhu and Zhang, 2010	Review volume, valence (10-item scale), and coefficient of variation	Sales	Video game	NPD GameSpot	Review volume has a positive influence on sales of games. Review valence has a positive influence on the sales only for less popular games. Review coefficient of variation has a negative influence on sales only for less popular games. Reviews (volume, rating, and variation) do not influence the sales of games without online capability.
Chen, Wang, and Xie, 2011	Review volume, valence (5-star scale), percentage of 1-star reviews, and percentage of 5-star reviews	Sales	Digital camera	Amazon	Review volume has a positive impact on sales. Review valence does not have an impact on sales. Percentage of 5-star reviews does not have an impact on sales. Percentage of 1-star reviews has a negative impact on sales.

Table 1.2 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Khare, Labrecque, and Asare, 2011	Review volume, valence (positive vs. negative), and consensus	Preference	Movie	Experiment	Interaction between review valence and volume: when review valence is negative, volume has negative impact on preference; when valence is positive, volume has positive impact on preference. Interaction among review valence, volume and consensus: when valence is positive and volume is high, low consensus decreases the preference; when valence is negative and volume is high, low consensus increases the preference; and when volume is low, consensus does not impact preference.
Moe and Trusov, 2011	Review number, valence, and variance	Subsequent preview posting and Sales	Bath, fragrance, and beauty products	A national retailer website	Increases in review valence encourage the subsequent posting of negative ratings, but discourage positive ratings. Increases in review variance negatively impact subsequent posting or extremely negative and extremely positive reviews. Increases in review volume increase all star level reviews. The magnitude of such impact is larger for negative ratings than for positive ratings. Baseline model: review valence, volume and variance all have positive impact on sales. Deviations from baseline model (caused by social dynamics): review valence directly (positively or negatively) affects sales. Variance and volume have indirect impact on sales.

Table 1.2 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Park, Gu, and Lee, 2011	Review valence and volume (as a moderator)	Sales	Digital camera	Amazon CNET Download.com	The impact of review valence from a specific website on a product positively interacts with its own volume and the volume of reviews for the same product from another website. The influence of review valence increases as its own volume increases, and decreases as the number of reviews on another website increases.
Zhou and Duan, 2011	Review valence (5-star scale)	Sales	Antivirus software, digital media player, download manager, and file compression	CNET Download.com	The increase in product variety weakens the impact of both positive and negative user reviews, and this weakening effect is more pronounced for popular products than on niche products.
Chen, Liu, and Zhang, 2012	Review absolute valence (unfavorable, favorable, andmixed) and relative valence (relatively negative, positive, and neutral)	Firm's financial value	Movie	IMDb, Yahoo! Movie, TNS media intelligence, 9 major US newspapers and 5 major entertainment publications	Relative valence influences firm value, but absolute valence does not, and the influence is greater during the prerelease period than the post-release period. For a given level of average valence, a larger number of earlier reviews may attract more investor attention and makes the deviation from it less impactful.

Table 1.2 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Sun, 2012	Review valence (5-star scale), volume, and variance	Sales	Book	Amazon.com Barnesandnoble.com	Review valence and volume have positive impacts on sales; standard deviation leads to relatively higher sales if and only if the valence is low.
Tirunillai and Tellis, 2012	Review valence (5-star scale), overall valence (positive vs. negative), and review number	Stock abnormal returns, idiosyncratc risk, and trading volume	Personal computing, cell phone, personal digital assistant or smartphone, footwear, toy, and data storage	Amazon, Epinions, and Yahoo! Shopping	Review volume has a significant positive impact on short-term and long-term stock returns. Number of negative reviews has a stronger impact on returns than positive reviews. Review valence does not impact stock returns. Review volume and the volume of negative reviews influence trading volume in both the short and long term. Negative reviews also positively influence firms' idiosyncratic risk. Off-line TV advertising increases review volume and decreases the number of negative reviews.

Reviews for Sellers

When reviewing the literature on the relationship between seller review and outcome, I found twenty-three studies that directly tested the consequences of seller reviews. Most of the products studied came from three categories: electronic products, such as digital cameras, laptops, MP3 players, and cell phones; collectable products, such as antique silverware, stamps, and gold coins; and entertainment products, such as books and DVDs. Lei (2011) chose a unique product to study: G-mail invitations. The nature of this product makes it possible to separate uncertainty related to sellers from uncertainty about the product condition. This product is only sold on the consumer-to-consumer market for a short period of time, and product condition does not vary. The only uncertainty related to the purchase is whether the seller will honestly deliver the product after the transaction (Lei, 2011).

Compared to research on product reviews, research on seller reviews focuses more on one particular market, eBay.com (seventeen out of the twenty three studies use data collected from eBay). One study (Wu and Ayala, 2012) tests the hypotheses with both experimental data and eBay transaction data. Since eBay reports several statistics, measurements of seller reputation show a little variation in the literature. The original eBay system allowed users to leave feedback for each other after each transaction, and eBay would summarize the number of positive, neutral, and negative reviews from unique users, along with a feedback score, or the number of positive reviews minus the number of negative reviews left by unique members. Weinber and Davis (2005) provide a snapshot of eBay's original review profile. eBay later added positive feedback percentage to user profiles. Positive feedback percentage is calculated as the number of positive reviews divided by the sum of positive and negative reviews left by unique members. A snapshot of an eBay review profile from2004 can be found in Zhang (2006). In 2007, eBay changed the calculation of positive feedback percentage by limiting the reviews included to those posted within a year instead of all reviews in a user's history, while maintaining the calculation of the feedback score. The cumulative counts of positive, neutral, and negative reviews throughout a user's history were no longer listed in user profiles. Figure 1.1 shows a snapshot of the newest review profile.

Many studies examine the impact of the feedback score, because this statistic combines review volume and review valence. While the feedback score has a consistent impact on sales and bidding participation, its impact on price is ambiguous. Feedback score has been shown to impact the price of auctions for G-mail invitations (Lei, 2011) and MP3 players (Sung and Liu, 2010) but not pennies (Lucking-Reiley et al., 2007) or magazines (Zhou et al., 2008). Even for the same product category, Obloj and Capron (2011) find that feedback score contributes to the price premium a seller can charge for cell phone auctions, but Huang et al. (2011) find no impact on auction price. The mixed results suggest that ratings that combine review volume and valence may not be sufficient for explaining consumers' preferences towards sellers.

As the impact of a single feedback score is unclear, many studies separate positive and negative reviews, using these two variables to indicate seller reputation independently. However, results are still mixed. The table below shows that separating positive and negative reviews still does not provide a clear picture of how seller reviews influence transaction outcomes, especially price. A negative review number does not always influence price (Ba and Paylou, 2002; Livingston, 2005), and the number of positive reviews can have a positive impact (Standifird, 2001; Ba and Paylou, 2002; Livingston, 2005; Houser and Wooders, 2006; Zhang, 2006; Reiley et al., 2007), no impact (Ba and Pavlou, 2002; Gilkeson and Reynolds, 2003) or even a negative impact (Gilkeson and Reynolds, 2003) on price.

Table 1.3 The Impact of Positive Reviews and Negative Reviews

Dependent Variable	Article	Number of Positive Reviews	Number of Negative Reviews
Price	Standifird, 2001	+	−
	Ba and Pavlou, 2002	+ for 13 products NS for 5 products	− for 2 products NS for 16 products
	Melnik and Alm, 2002		−
	Gilkeson and Reynolds, 2003	− or NS	
	Livingston, 2005	+	NS
	Houser and Wooders, 2006	+	−
	Zhang, 2006	+	NS
	Reiley et al., 2007	+	−
	Sung and Liu, 2010		NS
	Bockstedt and Goh, 2011	+	
Sales	Gilkeson and Reynolds, 2003	NS	
	Livingston, 2005	+	NS
Willingness to Bid	Melnik and Alm, 2002		−
	Livingston, 2005	+	−
	Park and Bradlow, 2005	NS	−

Zhou et al. (2008) compare different forms of review ratings provided by eBay and find that ratings that weight positive against negative reviews, such as review valence (the percentage of positive reviews), are more effective than feedback score in influencing auction price. Hence, to understand the role of reviews in consumers' decision-making processes, it is very important to look at the influence of review valence and review volume separately, as well as at the interaction between them (Khare et al., 2011; Park et al., 2012; Sun, 2012).

Table 1.4 provides a detailed summary of the results from the twenty-three studies.

MOTIVATION FOR MY RESEARCH

My research is motivated by the fact that research on online reviews has generated abundant information at the market, firm/product, and consumer levels; however, not enough studies incorporate consumer characteristics when investigating online reviews at the firm or

Table 1.4 Summary of Seller Review Outcomes

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Standifird, 2001	Number of positive reviews and of negative reviews	Final bidding price	3Com Palm Pilot V	eBay	Total number of positive reviews has a limited positive influence on bidding price. Total number of negative reviews has a negative influence on bidding price.
Ba and Pavlou, 2002	Positive review number and negative review number	Trust and price premium	Experiment and 18 products	eBay	Experiment study: Negative reviews have a stronger impact than positive reviews on buyer's trust in sellers. Trust mediates the relationship between reviews and price premiums. Product price moderates the relationship between trust and price premiums. Empirical study: Positive review number has positive impact on price premiums for 13 out of 18 products. Negative review number has negative impact on price premiums for only 2 of the 18 products. Product price only moderates the relationship between negative reviews and price premiums
Melnik and Alm, 2002	Feedback score and negative review number	Willingness to bid (WTB), and price	Gold coin	eBay	Feedback score has a significant positive impact on WTB and price. Negative review number has a significant negative impact on WTB and price.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Gilkeson and Reynolds, 2003	Number of positive reviews	Sales and closing price	Sterling silver flatware	eBay	Closing price is measured as the percentage of the average successful closing price. Number of positive reviews has no impact on auction success, either no or a negative impact on closing price.
Bruce, Haruvy, and Rao, 2004	Feedback score	Price	Laptop, PC, DVD, and book	eBay	Feedback score has a positive impact on price. The influence of feedback score is greater for low-price products than for high-price products.
Dewan and Hsu, 2004	Feedback score	Price and probability of sale	Stamp	eBay and MR	Prices are 10-15% lower on eBay than on MR. Feedback score has a statistically significant effect on auction price and probability of sale.
Livingston, 2005	Positive review number and percentage of negative reviews	Willingness to bid, sales, and price	Gold club	eBay	The number of positive reviews has a positive influence on bidders' willingness to bid, sales, and price, but the marginal effects diminish. Percentage of negative reviews has a negative influence on willingness to bid, but no influence on sales or price.
Park and Bradlow, 2005	Number of positive reviews and of negative reviews	Willingness to bid	Notebook	A Korean internet auction site	Number of positive reviews has no impact on willingness to bid. Number of negative reviews negatively influences willingness to bid.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Dewally and Ederington, 2006	STDDEV and negative review percentage	Final bidding price (the price of the item sold or the highest bid if not sold)	Collectable comic books	eBay	The mean negative percentage is 0.502%, and 59.4% of the sellers have no negative feedback. STDDEV measures how uncertainty about the negative portion by the standard deviation declines as the number of feedback increases; it has a negative impact on price, which means that total review number has a positive impact on price. Percentage of negative reviews has a negative impact on price.
Houser and Wooders, 2006	Number of positive, neutral, and negative reviews	Second highest bid plus shipping cost	Pentium III500 processors	eBay	Number of positive reviews has a positive impact on price. Number of neutral plus negative reviews has a negative impact on price.
Zeithammer, 2006	Feedback score	Highest and second highest bidding price	MP3 player	eBay	Feedback score has a positive impact on bidding prices.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Zhang, 2006	Number of positive reviews and number of negative reviews from buyer, seller, and both	Final bidding price and sales	Apple iPod MP3 player	eBay	Review from buyers: number of positive reviews positively influences the final bidding price; number of negative reviews negatively influences final bid and sales; no significant impact of the number of positive or negative buying reviews on final bids or sales. Total number of positive reviews positively influence final bid price, but total number of negative reviews does not influence final bid.
Chan, Kadiyali, and Park, 2007	Review valence (bidirectional); % of negative reviews	Willingness to pay	Notebook	A Korean internet auction site	Review valence has a positive impact on willingness to pay. Negative reviews do not have any impact on willingness to pay more than neutral reviews.
Reiley, Bryan, Prasad, Reeves, 2007	Feedback score, number of positive reviews, and number of negative reviews	Final bidding price	US Indian Head pennies	eBay	Feedback score does not influence price. Total number of positive reviews has a positive influence on price. Total number of negative reviews has a negative influence on price. The impact of negative reviews is larger than that of positive reviews.
Ghose, 2009	Review valence (5-star scale), review volume, % of positive reviews, % of negative reviews, price	Number of days it takes for product to be sold	Used laptop, digital camera, audio player, and PDA	Amazon Marketplace	Review valence has a positive influence on the time it takes to sell products. Review number has a positive influence. Percentage of positive reviews has a positive influence. Percentage of negative reviews has a negative influence.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Li, Srinivasan, and Sun, 2009	Feedback score	Willingness to bid, bidding amount, and entry and bidding time	Antique painting and silver plate	eBay	High feedback score encourages bid participation, decreases bidders’ bidding amounts, and encourages bidders to bid early. The impact of credibility of seller is stronger for bidders with more experience.
Zhou, Dresner, and Windel, 2009	Number of positive reviews, number of negative reviews, and feedback score	Final bid price	Digital camera	eBay	Direct counts (within the last 12 months) of positive and negative reviews significantly influence the final auction price. Feedback score and the difference between positive and negative review number within last 12 months do not significantly influence price. The effect of negative review number is larger than positive review number Review valence (positive percentage) has a significant influence on price.
Sung and Liu, 2010	Feedback score and negative review number	Price (winning bid plus shipping)	iPod shuffle MP3 Player	Yahoo! Taiwan	Feedback score has a positive impact on price. The impact of feedback score is significantly different across reputation quartiles; negative review number does not have impact on price.
Bockstedt and Goh, 2011	Review valence (bidirectional) and number of positive reviews	Price premiums	Nintendo Wii	eBay	Total number of positive reviews is significantly associated with higher price premiums. Review valence does not have a significant effect on price premiums because of a large concentration of high positive percentages.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Huang, Chen and Lu, 2011	Feedback score	Auction success and winning price	Nokia 8250	Yahoo!	Feedback score significantly affects auction success, but not auction price.
Lei, 2011	Feedback score and Feedback score related to selling gmail invitation	Sales and final bidding price	Gmail invitation	eBay	Feedback score has a positive impact on probability of sales and price. The squared feedback score has a negative impact on price. Feedback score related to selling Gmail invitation has no impact on price.
Obloj and Capron, 2011	Seller review difference	Price premium (difference in price)	New mobile phone	Polish internet auction site	Seller review difference is the difference in feedback scores between seller and competitor divided by the sum of feedback scores. The price premium a reputable seller can charge increases with the size of the reputation gap (the difference in reputation) between the seller and its matched competitor, but with a diminishing rate.

Table 1.4 Continued

Article	Independent Variable(s)	Dependent Variable(s)	Product(s)	Source(s)	Results
Wu and Ayala, 2012	Seller review valence (bidirectional) and volume	Willingness to pay (absolute and relative)	DVD set and iPod	Experiment and eBay	Experimental data results: Review volume has no impact on absolute willingness to pay, and it has a positive impact on relative willingness to pay for risk-averse and risk-neutral consumers, but no impact for risk-seeking consumers. Review valence has a positive impact on both absolute and relative willingness to pay for all consumers. Product price has a positive impact on absolute willingness to pay; it has a negative impact on relative willingness to pay for risk-neutral consumers, but no impact for risk-averse or risk-seeking consumers. Empirical data results: Review volume has no impact on absolute or relative willingness to pay for risk-averse and risk-seeking consumer, but a positive impact for risk-neutral consumers. Review valence has positive impacts on both absolute and relative willingness to pay for all consumers Product price has a positive impact on absolute willingness to pay for all consumers; it has a negative impact on relative willingness to pay for risk-averse consumers, no impact for risk-neutral consumers, and a positive impact for risk-seeking consumers.

market level. Especially with respect to review outcomes, a review of the literature indicates large inconsistencies in empirical results and conclusions. As mentioned by Wu and Ayala (2012), at the market or product level, variations cannot fully explain the inconsistency. The influence of online reviews must be understood from the consumer's standpoint, and that understanding should be incorporated into managing online reviews at the product and firm level.

To the best of my knowledge, Wu and Ayala (2012) is the first study that theorizes the influence of online reviews and consumer differences in price decisions, and investigates the impact of reviews at the consumer/individual level. They draw a theoretical framework from classical expected utility theory and incorporate seller's review volume and valence into consumers' judgment of risk level associated with the purchase. They propose that review volume and valence independently and directly impact a consumer's judgment of purchase risk, which influences the price she is willing to pay for the seller. Because consumers can have different risk attitudes, for example, risk averse, risk neutral, or risk seeking, reviews have different effects on willingness to pay. Review valence should always positively influence a consumer's willingness to pay, but the influence of review volume varies across consumer segments based on risk attitude. For risk-averse consumers, review volume has a positive impact on willingness to pay; for risk-neutral consumers, volume has no impact;and for risk-seeking consumers, volume has a negative impact.

Sun (2012) studies consumer heterogeneity from another perspective. She assumes that all consumers are risk neutral, but differ in their taste for the product. In contrast toWu and Ayala's (2012) research, she does not address how the difference in taste leads to different behaviors from consumers. In other words, she does not investigate the impact of heterogeneity at the consumer/individual level, but only uses the existence of heterogeneity to explain how review variances interact with valence in influencing product sales at the firm/product level. Although Sun (2012) does not investigate the impact of online reviews across heterogeneous consumer segments, she makes an observation at the aggregated level that cannot be explained by Wu and Ayala's (2012) framework.

Motivated by these studies, I develop a framework that will combine the strengths of both perspectives. First, consistent with the study by Wu and Ayala (2012), I incorporate consumer-level characteristics and look at different behaviors across consumers. Second, my framework can accommodate consumer behavior that leads to the interactions observed by Sun (2012).Third, my framework not only can account for the interactions between review valence and variance, but also can explain the three-way interactions between review valence, variance, and volume as documented by Khare et al. (2011). As a result, my framework predicts that online reviews can have opposite outcomes not only across consumers, but also within consumers.

Specifically, I focus on exploring the bidirectional review system and how it is used by consumers to shape their willingness to pay for sellers with different review profiles. The following considerations play a part in my focus.

First, online seller reviews are the most important online user reviews. Consumers can obtain information about products from other channels; for example, a consumer can visit local stores to check out the product and then purchase online. However, for sellers, most of the time consumers do not have comparable opportunities offline and online reviews become the main source of information. Wu et al. (2012) find that consumers perceive more uncertainty relating to online sellers than to products. The impact of online reviews should be more salient for sellers than for products, if not the same. So seller reviews are a good starting point for studying online reviews.

Second, bidirectional review is the most popular system used in online seller reviews. Each bidirectional review follows a Bernoulli distribution and a sample of reviews follows a binomial distribution (Wu and Ayala, 2012). Only volume and valence are important, because review variance is fully determined by the volume and valence. For a star-scale review system, review volume, valence, and variance are independent, and all three statistics are relevant when analyzing the characteristics of reviews. It makes sense to first analyze the impact of online reviews with respect to two variables, and then move to three variables.

Third, as mentioned above, studies of the relationship between seller reviews and sale price have found the most inconsistent results. I also find that studying the impact on price is more interesting because, as consumers, we can always make decisions on whether to purchase, either offline or online, but we do not have the freedom to make decisions on price for offline purchases. Online purchases provide a great opportunity for directly studying consumers' price decisions, or their willingness to pay, and can provide firms with insights that would be hard to obtain in offline settings.

My dissertation is organized as follows. In Essay Two, I describe the development of my theoretical framework and test its internal validity with an experimental study. In Essay Three, I test the external validity of my framework using a dataset of online transactions.

ESSAY TWO. ONLINE REVIEWS AND CONSUMERS' WILLINGNESS TO PAY: THEORETICAL FRAMEWORK AND AN EXPERIMENTAL INVESTIGATION

INTRODUCTION

As discussed in Essay One, my research will explore how a bidirectional review for a seller influences consumers' purchase decisions. More specifically, my study will focus on how a seller's review number (volume) and percentage of positive reviews (valence) influence consumers' willingness to pay.

A review of the literature reveals that empirical studies have generated mixed results concerning the relationship between seller review statistics and the price a seller can charge. A closer look at these studies shows that the measurements related to review volume, such as feedback scores, number of positive reviews, and number of negative reviews, all have inconsistent influences on price. On the other hand, review valence has shown a relatively consistent influence. As shown in Table 2.1,only one study finds that review valence has no significant impact on price; however, Bockstedt and Goh (2011) explain that this result may be due to the large concentration of high review valence in the dataset. Similar findings are documented by Wu and Ayala (2012), who find that review valence consistently influences price, but the impact of review volume is ambiguous.

Table 2.1 The Impact of Review Valence on Price

Article	% of Positive Reviews (Valence)	% of Negative Reviews
Livingston, 2005		NS
Chan et al., 2007	+	
Ghose, 2009	+	−
Bockstedt and Goh, 2011	NS	
Wu and Ayala, 2012	+	

Several studies try to provide theoretical explanations for the mixed results observed in empirical studies. The authors break down the aggregate data to directly examine at segment or even the individual level for possible explanations, postulating a relationship between seller reviews and consumers' willingness to pay. They propose that heterogeneity across consumers may explain the ambiguous relationship between online seller reviews, especially review volume, and the price of the product.

Kalyanam and McIntyre (2001) assume that, although all consumers are risk averse, they may differ in degree, and find that review volume and valence have an impact on price only for those consumers who are highly risk averse. Wu and Ayala (2012) and Wu et al. (2012) relax such a restriction and assume that consumers can be risk averse, risk neutral, or even risk seeking. If a consumer is risk averse, she will always prefer sellers with a large review volume; if a consumer is risk seeking, she will always prefer sellers with a small review volume; and if a consumer is risk neutral, she basically does not care about review volume. The common theme among these studies is that they use classic expected utility theory as a framework. Under such a framework, the preference towards risk or uncertainty (hence preference towards review volume) can vary across individuals but should be consistent within an individual. However, studies in other contexts such as insurance, warranties (Hogarth and Kunreuther, 1992), and financial assets (Sarin and Weber, 1993) have found that consumers change their preferences towards uncertainty depending on the probability of obtaining an outcome. I am interested in exploring whether such a change in preference will also occur when the purchase is made from online sellers. Thus, I hope to extend previous studies and provide further explanations for the impact of online seller reviews on willingness to pay by proposing that differences in preference towards uncertainty (hence towards review volume) not only exists across individuals, but also exists within an individual (at least for some consumers).

In the rest of Essay Two, I discuss the theoretical frameworks that model people's preference towards uncertainty, state specific hypotheses, present an experimental study, and, finally, propose an approach for testing the external validity of my framework.

THEORETICAL FRAMEWORK

Hereafter, I use the word "prospect" to represent a gamble-like problem. The remainder of this essay focuses on a prospect with two outcomes: obtaining outcome x if event E happens and outcome 0 if E does not. Let p be the true probability of event E, and a person will make decisions based on her assessment of the overall value of the prospect ($x, p; 0, 1-p$). There is a risk associated with outcome x when p is not 0 or 1;that is, we are not sure to obtain outcome x or 0. Many studies extend the concept of uncertainty using a more general framework in which not only is the outcome uncertain, the probability of each outcome is also ambiguous. Motivated by observations first documented by Ellsberg (1961), researchers have focused their efforts on theorizing decision rules that account for uncertainty generated partly by the risk of an unsure outcome and partly by the ambiguity concerning the probability of each possible outcome (Kahn and Sarin, 1988; Hoarth and Einhorn, 1990; Tversky and Kahneman, 1992; Fox and Tversky, 1995; Kilka and Weber, 2001).

Modeling Decisions under Uncertainty

The framework of expected utility theory. Expected utility theory holds that the over-all value of the prospect is

$$V(x,p;0,1-p) = EU[p \times x + (1-p) \times 0] = p \times U(x) + (1-p) \times U(0)$$

where $EU(\bullet)$ is the expected utility function and $U(\bullet)$ is the utility function with $U(0)=0$.

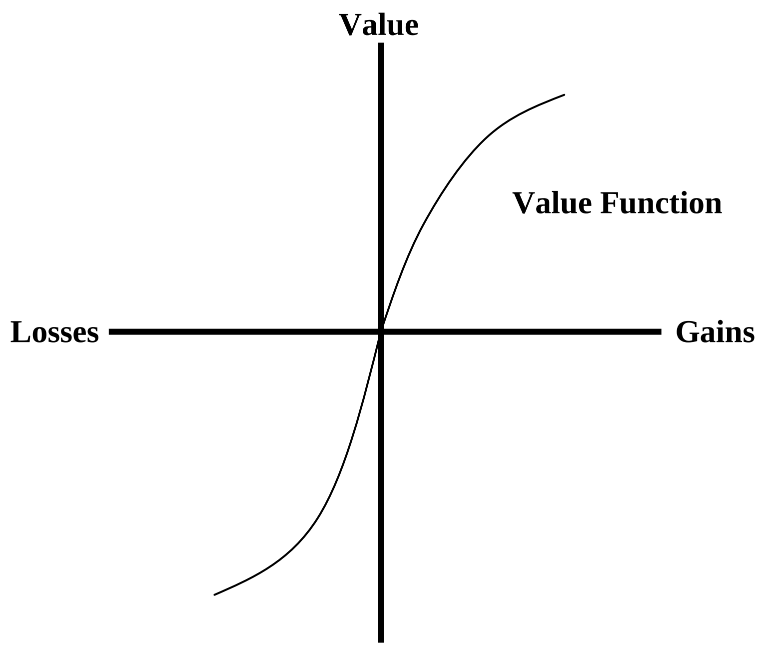
Generally speaking, expected utility theory assumes that a person's rational decision-making process follows certain assumptions (named axioms), and that under these assumptions we can obtain numeric measures of her utility assessments of a prospect's outcomes. A person will make decisions based on the utilities and probabilities of prospect's outcomes (Weber and Camerer, 1987; Hastie and Dawes, 2001).

The framework of prospect theory (PT) and cumulative prospect theory (CPT). Human behaviors that violate the assumptions of expected utility theory motivated the development of prospect theory (Kahneman and Tversky, 1979). To accommodate these behavioral patterns, prospect theory proposes separating the value and weighting functions when judging the overall value of a prospect.

$$V(x,p;0,1-p) = v(x)w(p) + \ v(0)w(1-p) = v(x)w(p)$$

where $v(\bullet)$ is the value function of an outcome and $v(0) = 0$, and $w(\bullet)$ is the weighting function for the stated probability.

<u>The properties of value function $v(\bullet)$.</u> Prospect theory holds that the value of an outcome conforms to a concept of "changes in wealth or welfare."A person does not measure the value of an outcome based on its final state, but rather on the difference between the final state and the current state of that person. Thus, the value function involves two aspects:"the asset position that serves as reference point and the magnitude of the change (positive or negative) from that reference point"(Kahneman and Tversky, 1979 page 277).Specifically, the value function is defined as generally concave for gains-outcomes' final states positively deviate from the reference point and convex for losses-outcomes' final states negatively deviate from the reference point, and is steeper for losses than for gains. The shape of the value function is shown below.



Source: Kahneman and Tversky, 1979

Figure 2.1 Value Function of Prospect Theory

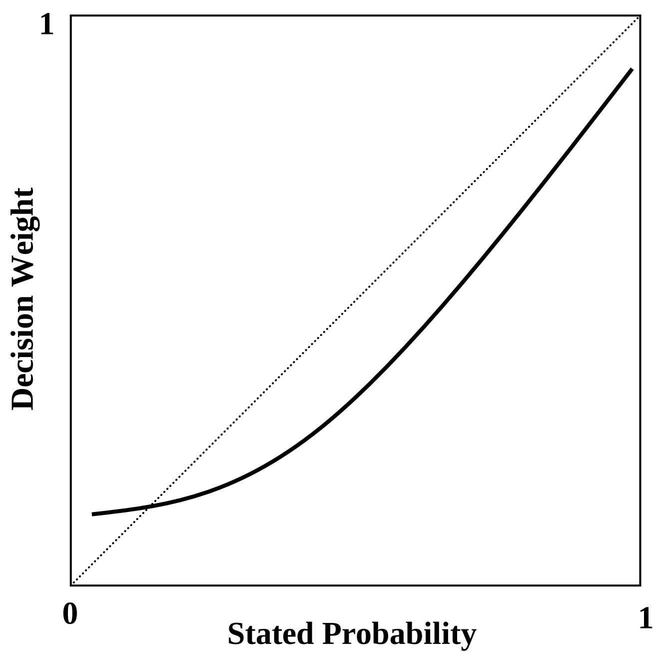
Since the value function is different for gains than for losses, a separation of decisions under gains and under losses becomes necessary. This separation can be achieved using gain- and loss-framing. In gain-framing, information is presented as a positive outcome as compared to a person's current state, with the associated probabilities, while in loss-framing, information is presented as a negative outcome. Examples of gain-framing and loss-framing are shown below.

Table 2.2 Examples of Decision Framings

Decision Scenario	A Gain Framing	A Loss Framing
Surgery (Compare to Radiation Therapy)	Of 100 people having surgery 90 live through the post-operative period, 68 are alive at the end of the first year and 34 are alive at the end of five years.	Of 100 people having surgery 10 die during surgery or the post-operative period, 32 die by the end of the first year and 66 die by the end of five years.
Gambles (Compare to A Sure Gain/ Loss)	25% chance to gain $1000 and 75% chance to gain nothing	75% chance to lose $1000 and 25% chance to lose nothing

Source: Tversky and Kahneman, 1986

The properties of weighting function $w(\bullet)$. Prospect theory transforms stated probabilities into decision weights that measure "the impact of events on the desirability of prospects" (Kahneman and Tversky, 1979 page 280). The weighting function is an increasing function of stated probability with $w(0) = 0$ and $w(1) = 1$. The slope of the weighting function measures the sensitivity of a person's preference towards the change in probability. Prospect theory describes several properties of weighting function: (1) overweighting: people tend to overweight very low probabilities, (2) subcertainty: the sum of weights associated with complementary events is generally less than the weight associated with the certain event, and (3) subproportionality: for a fixed ratio of probabilities, the ratio of corresponding decision weights is closer to unity when the probabilities are low. Based on these properties, a weighting function should have the shape shown in Figure 2.2.



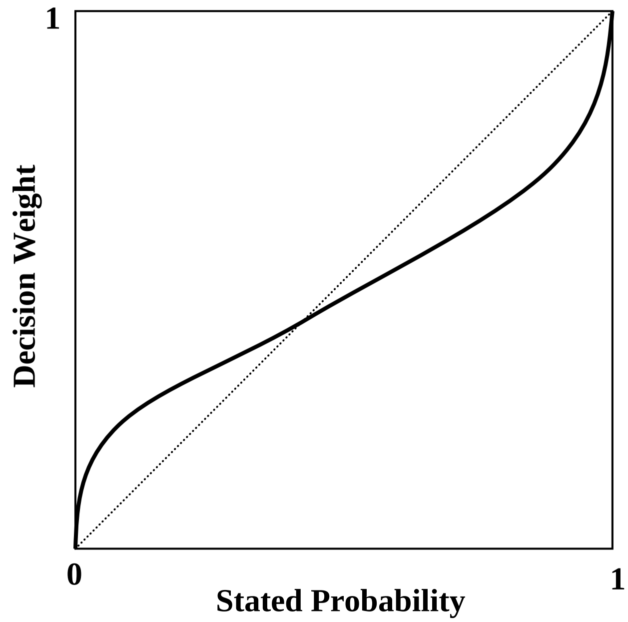
Source: Kahneman and Tversky, 1979

Figure 2.2 Weighting Function of Prospect Theory

In the more general uncertainty framework, additional uncertainty results from missing or ambiguous information about outcome probability *p*; hence, researchers usually use the weighting function to account for this type of uncertainty. In an update of prospect theory, Tversky and Kahneman (1992) point out that the weighting function in the original theory does not always satisfy stochastic dominance, cannot account for situations where outcome probabilities are unclear, and cannot be well extended to prospects with a large number of outcomes. To overcome these limitations, they re-frame the concept of prospect theory using rank-dependent utility theory, naming it cumulative prospect theory. Cumulative prospect theory allows for separate weighting functions for gains and losses. The decision weight associated with an outcome is interpreted as a marginal contribution of the outcome. Specifically, the decision weight associated with a positive outcome x_i is the difference between the capacities of events with the outcome that is “at least as good as x_i” and of events with the outcome that is “strictly better than x_i”(Tversky and Kahneman, 1992 page 301)The decision weight associated with a negative outcome is the difference between the capabilities of the events with the outcome that is “at least as bad as x_i” and of events with the outcome that is “strictly worse than x_i”(Tversky and Kahneman, 1992 page 301). However, when a prospect has two non-mixed outcomes, both positive or both negative, cumulative prospect theory and the original prospect theory yield the same prediction, because, under these conditions, original prospect theory is rank dependent.

Cumulative prospect theory proposes that the weighting function should satisfy both lower subadditivity, “the impact of an event A is greater when it is added to a null event than when it is added to some nonnull event B” (Tversky and Fox, 1995 page 270), and upper subadditivity, “the impact of an event A is greater when it is subtracted from the certain event than when it is subtracted from some uncertain event $A \cup B$”(Tversky and Fox, 1995 page 271).

These characteristics lead to a reversely *S*-shaped weighting function, which holds that people tend to overweight small probabilities and underweight large probabilities.



Source: Tversky and Kahneman, 1992

Figure 2.3 Weighting Function of Cumulative Prospect Theory

Tversky and Fox (1995) test the characteristics of the weighting function in the more general uncertainty situations in which ambiguity about outcome probability exists, and they find that both lower and upper subadditivities apply to the more general uncertainty, and such effects are amplified when outcome probabilities are unclear. The smooth function used to fit the weighting function in the cumulative prospect theory is shown below. The same form of the weighting function also has been tested by many other studies (Wu and Gonzalez, 1996; Prelec, 1998; Schimdt et al., 2008).

$$w(p') = \frac{p'^{\gamma}}{(p'^{\gamma} + (1 - p')^{\gamma})^{1/\gamma}}$$

where p' is the stated probability, and γ is the parameter that influences the shape of the weighting function and can be set to different values for gains versus losses.

Other frameworks for decisions under uncertainty. Kahn and Sarin (1988) extend subjective utility theory (which is an extension of classic expected utility theory) and propose a decision model that depends on the entire distribution of p, assuming p is a random variable.

Let $\varphi(p)$ denote the density of the random variable p and $\bar{p}$ is the average of p. Then the over-all value of the prospect $(x, p; 0, 1-p)$ is:

$$V(x, p; 0, 1-p) = w(p)u(x);\ w(p) = \bar{p} + \int_{p=0}^{1} (p - \bar{p}) e^{\frac{[-\lambda(p-\bar{p})]}{\sigma}} \varphi(p) dp$$

$\sigma = \sqrt{\int_{p=0}^{1} (p-\bar{p})^2}\, \varphi(p)dp$ is the standard deviation of the random variable p, and λ is a person's attitude towards uncertainty about probabilities. Using a first-order Taylor approximation of $e^{\frac{[-\lambda(p-\bar{p})]}{\sigma}}$, the weighting function $w(p)$ can be expressed by the first-order approximation

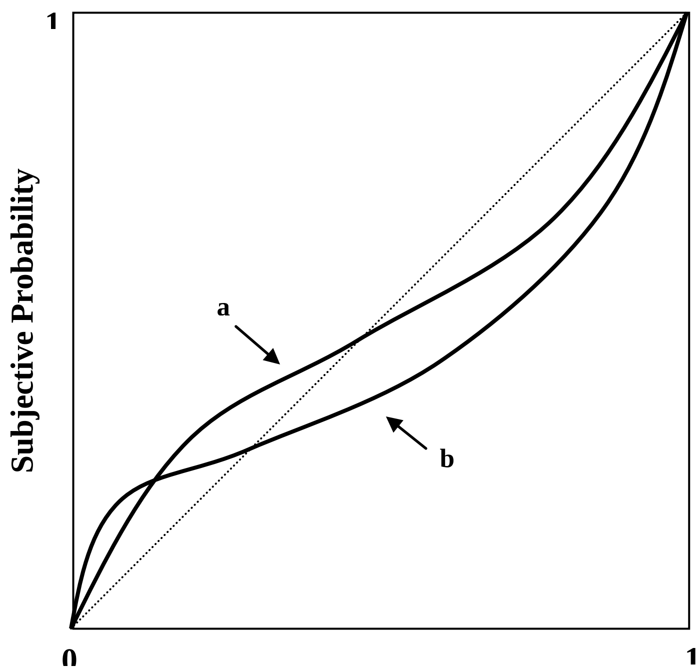
$$w(p) = \bar{p} - \lambda\sigma$$

The weighting function states that the subjective probability of event E is deviated from average probability $\bar{p}$ and that the deviation is related to the standard deviation of the random variable p.

The framework proposed by Einhorn and Hogarth (Einhorn and Hogarth, 1985, 1986; Hogarth and Einhorn, 1990) is built on an idea similar to prospect theory. The authors propose that a person may take an anchoring-and adjustment strategy. When the true probability p of event E (hence the outcome) is unknown, a person can start with an anchor, such as the stated probability p', and then make either an upwards or downwards adjustment based on the level of probability, amount of uncertainty perceived, and the person's attitude towards uncertainty about probabilities. The weighting function is expressed as follows:

$$w(p') = p' + \theta(1 - p' - p'^{\beta})$$

where θ ($0 \leq \theta \leq 1$) is the absolute size of adjustment and β is the attitude towards uncertainty about probabilities. β> 1 implies that a person places more weight on probabilities that are larger than p' and that$w(p')$ is larger than $p'$for most of the range of p'.In contrast, β <1 implies that a person places more weight on probabilities that are smaller than $p'$and that $w(p')$ is smaller than p' for most of the range.



Subjective Probability

Pavoff Size: a<b

Source: Einhorn and Hogarth, 1985

Figure 2.4 Weighting Function of Einhorn and Hogarth's Model

Many researchers have proposed other smooth models to describe a reversely *S*-shaped weighting function. For example, Prelec (1998) proposes an exponential function with either one parameter or two parameters; Gonzalez and Wu (1999) propose a nonparametric estimation of weighting function at individual level, and by using this method they find that a two-parameter "linear in log odds model" weighting function was superior than one-parameter model in the domain of gains, one of which represents how a subject discriminates probability and the other of which measures how attractive the gambling is.

The difference between expected utility theory and the behavioral theories discussed above is that under expected utility theory, the overall value of a prospect is determined by the true probability *p*, which is assessed by the stated sample probability p' through a mean-variance model. Behavior theories argue that when information about *p*is missing, people in general experience extra uncertainty resulted from the ambiguity of probability. They propose nonlinear transformations from the stated probability p' to a subjective weight, either a decision weight that measures the desirability of event *E* (Tversky and Kahneman, 1992; Gonzalez and Wu, 1996; Prelec, 1998) or a subjective probability that measures subjective likelihood of event *E* (Einhorn and Hogarth 1985; Kahn and Sarin, 1988). Depending on the shape of the transformation function, a person may overweight or underweight probabilities based on the level of the stated probability. The magnitude of overweighting/underweighting is related to various factors. If the stated probability p' comes from a sample used to estimate the true probability of an outcome, then the size of the sample, the source credibility of the sample, and the degree of agreement or disagreement among the sources should influence the magnitude of the adjustment (Einhorn and Hogarth, 1985, 1986; Camerer and Weber, 1992). The larger the sample size, the higher the source credibility, and the smaller the disagreement among sources, the smaller the magnitude of adjustment and subjective probabilities will approach the stated probabilities.

Preference towards Uncertainty

The framework of expected utility theory. According to expected utility theory, people's preferences towards risk can be different. When comparing two prospects, prospect A ($x, p; 0, 1\text{-}p$) and prospect B ($xp, 1; 0, 0$),it is obvious that B is the certainty equivalent of prospect A. However, not all people see them as equal. For people who place more value on B than A (hence prefer B to A), the suggestion is that they value a certain prospect more than an uncertain one, and therefore that these people are risk averse. People who place more value on A than B (hence prefer A to B) are risk seeking. Finally, people who are indifferent to the prospects A and B are risk neutral. This concept, called risk attitude, establishes the difference between individuals in terms of their preferences towards uncertainty. Mathematically, a person's risk attitude can be determined by the shape of her expected utility function (determined by x and p together), with a concave expected utility function representing risk averse, a convex expected utility function representing risk seeking, and a linear expected utility function representing risk neutral. Risk attitude can be interpreted as a kind of personality, which is consistent under a specific context. So if a person is risk averse, she will always prefer the certainty equivalent regardless of whether p is large or small.

The framework of prospect theory and cumulative prospect theory. According to prospect theory and cumulative prospect theory, a person's preference towards uncertainty is determined jointly by the value function and the weighting function. In general, a reversely *S*-shaped weighting function plus a concave-shaped value function for gains leads to uncertainty-seeking behavior for gaining a prize of small probabilities and uncertainty-averse behavior for gaining a prize of large probabilities. The same shape of weighting function plus a convex-shaped value function for losses leads to uncertainty-averse behavior for losing at small probabilities and risk-seeking behavior for losing at large probabilities.

Other frameworks for decisions under uncertainty. Einhorn and Hogarth (1985,1986) propose a pattern similar to that found in prospect theory. They assume that people are generally defensive pessimistic about gaining something: they will overweight small probabilities and underweight moderate to large probabilities. Hence, they display risk-averse behavior for gaining at small probabilities and risk-seeking behavior for gaining at moderate and large probabilities.

Kahn and Sarin (1988) do not directly describe the behavior of the majority under uncertainty; however, they advocate the idea that a person's preference towards uncertainty can vary with $\bar{p}$.In other words, a person's preference towards uncertainty can depend on the expected probability that event *E* will happen. A simple way of incorporating such a variation of preference into their model is to allow a person's attitude towards uncertainty λ to be a linear function of $\bar{p}$. As a result, the first-order approximation of their model can be rewritten as follows:

$$w(p) = \bar{p} - (a + b\bar{p})\sigma$$

Expected utility theory acknowledges the differences across individuals in terms of preference towards uncertainty; however, that preference should be consistent within an individual and independent of the probability of obtaining an outcome. In contrast, behavior theories propose that even within an individual, a change in preference towards uncertainty may occur, depending on the level of stated probability p'.

HYPOTHESES DEVELOPMENT

In this section, I develop the proposition regarding consumer heterogeneity, and propose the influences of review volume and valence for different consumers.

Online Purchase Decision: Willingness to Pay (*WTP*)

Assume that the reference point of a consumer before purchasing a product with monetary value *V* from an online seller is 0. The purchase can be simplified to a two-outcome prospect: a consumer is either being satisfied by the seller or not, where the value of being satisfied is gaining *V* and the value of being not satisfied is 0. Let *p* denote the true probability that a consumer will be satisfied by the online seller for the transaction. The purchase decision in terms of willingness to pay can be represented as how much the consumer is willing to pay to purchase the prospect (*V, p; 0, 1-p*). Assuming that when determining the price of a prospect, people tend to evaluate the outcome (*V, p; 0, 1-p*) and cost (*-WTP, 1*) separately (Kahneman and Tvresky, 1979), the maximum willingness to pay (WTP) will be determined as

$$v(V, p; 0, 1\text{-}p) + v(\text{-}WTP) = 0$$

Seller review information is presented as the probability of gaining the product and being satisfied by the seller, so the framing of a bidirectional seller review is a gaining framing. The basic idea behind seller reviews is that a consumer does not know the true probability *p* that she will be satisfied by the seller and can only use a sample to estimate *p*. Previous customers who provide reviews about a seller form a sample of the population (all customers of that seller). A bidirectional review system entails that the sample follows a binomial distribution; the review volume *N* is the size of the sample; the review valence p', the percentage of positive reviews, is the sample mean; and the estimator of *p*,$\frac{p\prime(1-p\prime)}{N}$,is the variance of the sample meanp'(Wu and Gaytan, 2010). According to prospect theory and cumulative prospect theory, the price a consumer is willing to pay is then determined by the subjective value of the prospect (*V*, p'*; 0, 1-*p') as perceived by that consumer.

$$WTP = -v^{-1}[-v(V)w(p', N) - v(0)w(1-p', N)] = -v^{-1}[-v(V)w(p', N)]$$

where $v(V) > 0, v(0) = 0$, and *w(*p'*)* is an increasing function of p'

As the main focus of the current research is the weighting function, which previous studies have theorized is the source of reversed preference towards uncertainty for gains with different probabilities, I will set the value *V* at a constant level and examine a consumer's preference towards uncertainty using the weighting function *w(*p'*, N).*

Proposition: The Shape of Weighting Function $w(p', N)$

As discussed earlier, previous research (Wu and Ayala, 2012; Wu et al., 2012) developed using the framework of expected utility theory assumes that uncertainty preference only differs across individuals. While behaviorists argue that people tend to have reverse preferences towards uncertainty for gains of small probability and for gains of large probability, List (2004) finds that in the marketplace, prospect theory explains the behavior of inexperienced consumers well. However, consumers with greater market experience tend to conform to the predictions of classic expected utility theory. In market like eBay, where variety exists among consumers, it is reasonable to expect to observe the behavioral patterns predicted by both frameworks. I extend the previous research by allowing the differences in preference towards uncertainty to exist not only across individuals, but also within an individual (at least for some consumers). So, the assumption is that there is heterogeneity of consumers in the shapes of the weighting function. First, there are consumers who consistently overweight or underweight all probabilities. For these consumers, their preference towards uncertainty can be described by the risk attitude of expected utility theory. Specifically, if a consumer consistently overweights all probabilities (a concave-shaped weighting function), then she is a risk-seeking person; if a consumer consistently underweights all probabilities (a convex-shaped weighting function), then she is a risk-averse person; and if a consumer neither overweights nor underweights any probability (a linear weighting function), then she is a risk-neutral person. Second, there are consumers who do not consistently overweight or underweight probabilities. As discussed before, a reversely *S*-shaped weighting function has been proposed under the assumption that people are generally defensive pessimistic about gains (Einhorn and Hogarth, 1985,1986). I relax this restriction, allowing consumers to have a weighting function that is *S*-shaped. Consumers who underweight small probabilities and overweight large probabilities have weighting functions with an *S* shape, and consumers who overweight small probabilities and underweight large probabilities have weighting functions with a reversed *S* shape. For consumers with either an *S*-shaped or reversely *S*-shaped weighting function, there exists a cross-over point (Einhorn and Hogarth, 1985, 1986) where $w(p', N) = p'$.

Proposition. A consumer's weighting function is

a. concave if she overweights all probabilities.
b. convex if she underweights all probabilities.
c. linear if she neither overweights nor underweights any probability.
d. *S*-shaped if she underweights small probabilities and overweights large probabilities.
e. reversely *S*-shaped if she overweights small probabilities and underweights large probabilities.

Hypothesis 1: The Impact of Seller Review (SR) Valence (*p'*) on Willingness to Pay

Review valence is the stated probability obtained from a sample and is used to estimate the true probability of obtaining outcome *v(V).* Consistent with the underlying assumption of previous behavioral studies that the weighting function should be an increasing function of the stated probability, the weighting function $w(p', N)$ should be an increasing function of review valence p'. Hence the impact of review valence p' on WTP should be positive for all consumers.

H1. For a seller with a higher SR valence (p′), a consumer is willing to pay a higher price regardless of the shape of the consumer's weighting function.

Hypothesis 2: The Impact of Seller Review Volume (*N*) on Willingness to Pay

Review volume *N* is the size of the sample from which review valence p' is obtained. The larger *N* is, the smaller the magnitude of overweighting/underweighting. As a result, the impact of review volume on WTP depends on whether a consumer overweights or underweights the review valence. For consumers who consistently overweight or underweight probabilities, the impact of review volume on WTP is consistent, either negative, positive, or insignificant. For consumers who do not consistently overweight/underweight, the impact of review volume *N* is determined by the shape of a consumer's weighting function and by the level of review valence, specifically, whether valence is below or above the cross-over point.

H2. For a seller with a higher SR volume (*N*), a consumer is willing to pay

a. a lower price if the consumer has a concave-shaped weighting function.

b. a higher price if the consumer has a convex-shaped weighting function.

c. an equal price if the consumer has a linear-shaped weighting function.

d_1. a higher price if the consumer has an *S*-shaped weighting function and the SR valence (p′) is below the cross-over point.

d_2. a lower price if the consumer has an *S*-shaped weighting function and the SR valence (p′) is above the cross-over point.

e_1. a lower price if the consumer has a reversely *S*-shaped weighting function and the SR valence (p′) is below the cross-over point.

e_2. a higher price if the consumer has a reversely *S*-shaped weighting function and the SR valence (p′) is above the cross-over point.

The overall conceptual framework is shown below.

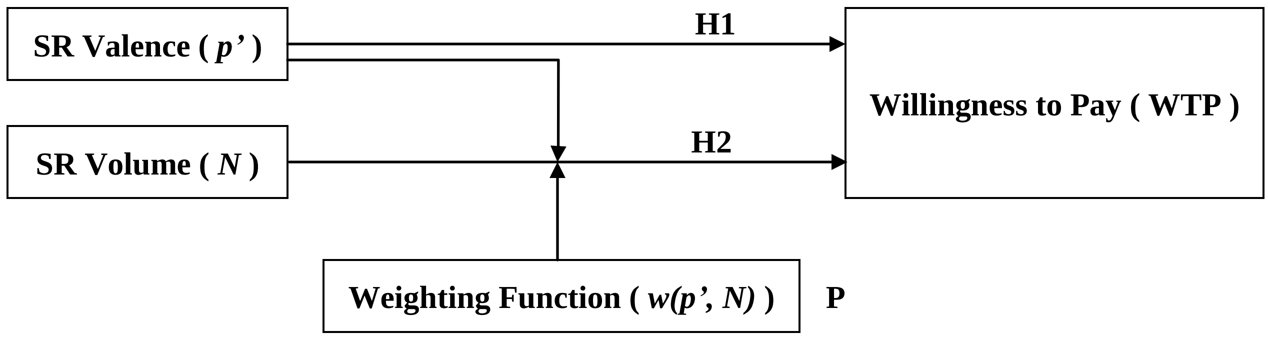


Figure 2.5 Conceptual Framework

AN EXPERIMENTAL STUDY

Study Design

Subjects are asked to consider a purchase scenario in which they are about to purchase a 42" LCD TV on a website. The TV is sold at local retail stores for $800. On the website, there are multiple sellers selling the new TV and the website provides reviews for each seller. Seller review has a bidirectional format, as shown below.

The review volume has three levels: 20, 50, and 200, and the review valence has eleven levels: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%. Each subject was provided with 33 seller profiles having different combinations of levels of volume and valence. Four versions of the survey were developed to counterbalance the order in which volumes and valences were shown to the subjects.

Data Collection Procedure

One-hundred forty-three business-school students at a southern public university were recruited for the study. Each subject was randomly assigned to one of the four versions of the survey. Subjects were asked to provide the maximum price they were willing to pay each seller for the product; the lowest price they could pay was $0. Three seller profiles appeared twice in the survey to test the inner reliability of the answers provided by each subject. These profiles were (20, 50%), (50, 50%), and (200, 50%), and appeared in the middle and then the end of the survey.

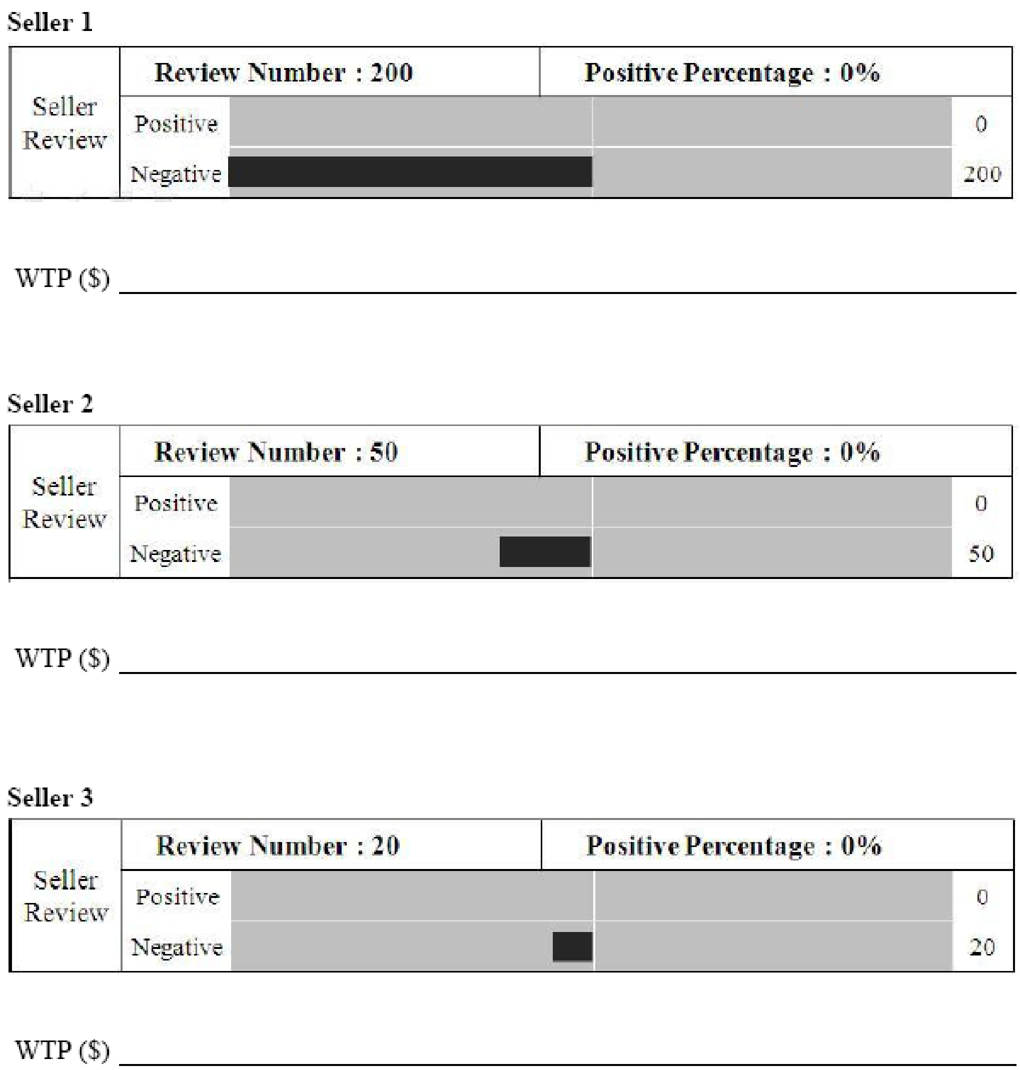


Figure 2.6 Experiment Study Design Snapshot

Analyses

Internal reliability. Internal reliability was assessed on the repeated seller profiles using the Pearson Correlation test. Subjects with a Pearson Correlation value below 0.8 were removed from the dataset. For subjects who could not be tested using Pearson Correlation because they had the same reported WTP for different sellers, the pattern of reported WTP based on review volume was observed. Subjects with dramatic pattern changes, such as a reversed preference on review volume between the test and re-test sets, were removed from the dataset.

Assess the shape of weighting function. The following model was used to estimate the weighting function for each subject:

$$w_i(p', N) = c_i + p' - (a_i + b_i p')\frac{p'(1-p')}{N}$$

where *a*, *b*, and *c* are parameters, and *i* represents the i^{th} individual

Because product value is set at a constant level, I use an intercept, *c*, to capture the deviation of the subjective product value from the objective product price.

The weighting function combines ideas from Kahn and Sarin (1988) and Einhorn and Hogarth (1985, 1986), adopting the form used by Kahn and Sarin (1988). Both studies state that the sample variance of random variable *p'* will positively impact the magnitude of uncertainty. Einhorn and Hogarth (1985, 1986) also propose that magnitude is negatively associated with factors such as sample size and source credibility. As source credibility is not the focus of my dissertation, my weighting function model is only a function of sample variance $p'(1-p')$ and sample size *N*. The term $(a + bp')$ describesa subject's attitude towards uncertainty as a function of *p'*, which allows the attitude to change at different levels of *p'*.

A slight modification was made without changing the properties of the function shapes. I use sample variance instead of standard deviation, used by Kahn and Sarin (1988). I chose this model specification for several reasons. First, the model directly incorporates the variables in which I am interested into the estimating weighting function. Second, the model can accommodate all five types of shapes. Lastly, using variance, the shape of the weighting function can be directly determined by the estimates of parameters *a* and *b*. Subjects with the same shape of weighting function were then grouped together. For the *S*-shaped and reversely *S*-shaped weighting functions, the cross-over point is $-\frac{a_i}{b_i}$. Table 2.3 below shows how to use estimates of *a* and *b* to determine the shape of the weighting function for each subject.

Table 2.3 Estimation of the Shape of Weighting Function

Group	*w*(*p*) Shape	Description	Parameters
1	Concave	Overweight all probabilities	b=0 and a<0, *b*>0 and *a/b*≤ -1, *b*<0 and *a/b*≥ 0
2	Convex	Underweight all probabilities	*b*=0 and *a*>0, *b*>0 and *a/b*≥ 0, *b*<0 and *a/b*≤ -1
3	Linear	Neither underweight nor overweight probabilities	*a* = 0 and *b* = 0
4	*S*-Shaped	Underweight small probabilities Overweight large probabilities	*b* < 0 and -1< *a/b* < 0
5	Reversely *S*-Shaped	Overweight small probabilities Underweight large small probabilities	*b* > 0 and -1 < *a/b* < 0

Assess the impact of seller review valence and volume. The impact of SR valence and SR volume on WTP were assessed at the group level. For each group, formed by the shape of the weighting function, a linear regression was performed to test the impacts of valence and volume. I used a linear-log function because it was used in previous empirical research to test the relationship between seller reputation and price (Ba and Pavlou, 2002; Melnik and Alm, 2002; Lucking-Reiley et al., 2007; Huang et al., 2011).The main reason for using a linear-log as opposed to a linear function is that it can capture the diminishing return of reputation on price as the seller reputation increases (Livingston, 2005; Obloj and Capron, 2011). For the *S*-shaped and reversely *S*-shaped weighting function groups, separate linear regressions were used to fit the data that fell below or above the cross-over point.

$$Log(WTP_{ijk}) = \alpha_{ijk} + \beta_{pjk}Log(p') + \beta_{Njk}Log(N) + \varepsilon_{ijk}$$

where i identifies the i^{th} individual, j represents the j^{th} group, and k denotes below or above cross-over point.

Results

Internal reliability. Twenty-eight of the one-hundred forty-three students did not pass the internal reliability test and hence were removed from the original dataset. Examples of answers from those subjects are shown in Figure 2.7.

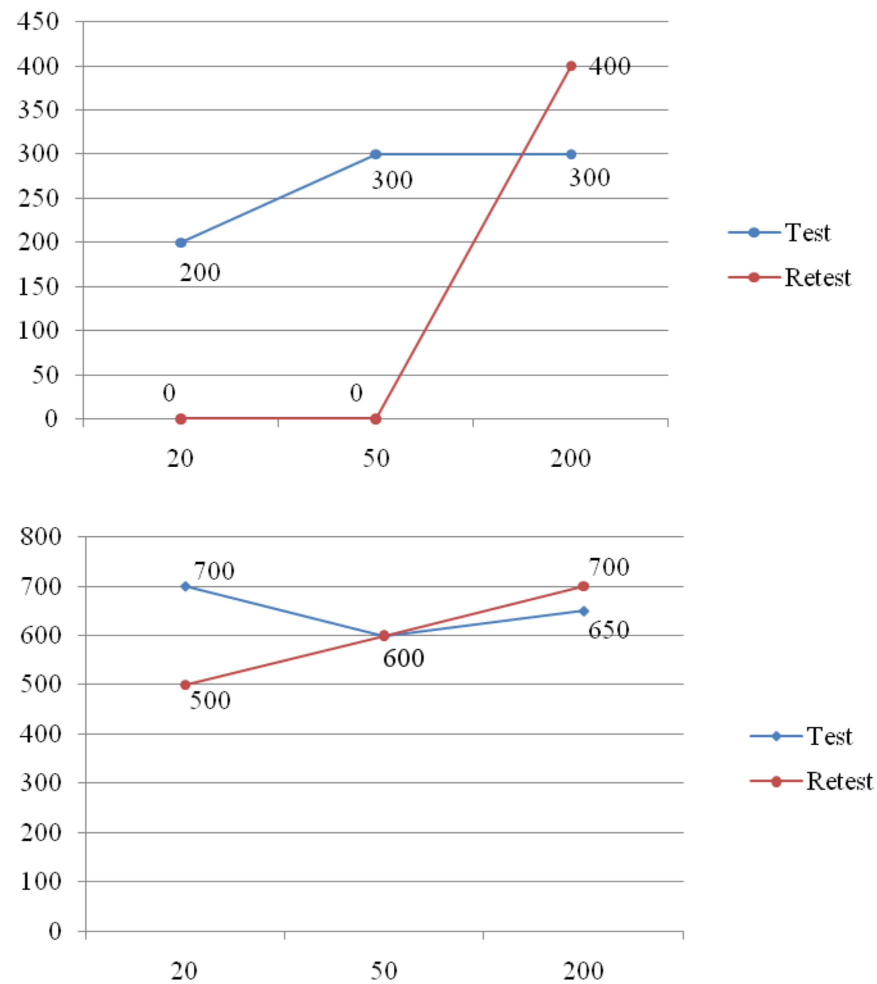


Figure 2.7 Examples of Subjects Removed from the Data

The shapes of weighting function. All five groups of weighting function shapes were identified, supporting the proposition. For those subjects who consistently overweight/underweight all probabilities, 10 subjects have concave-shaped weighting functions corresponding to a risk-seeking attitude, 7 have convex-shaped weighting functions corresponding to a risk-averse attitude, and 22 subjects have linear-shaped weighting functions corresponding to a risk-neutral attitude. For those subjects who do not consistently overweight/underweight probabilities, 25 have *S*-shaped weighting functions and 51 have reversely *S*-shaped weighting functions. The plots of weighting functions by groups are provided in Figure 2.8.

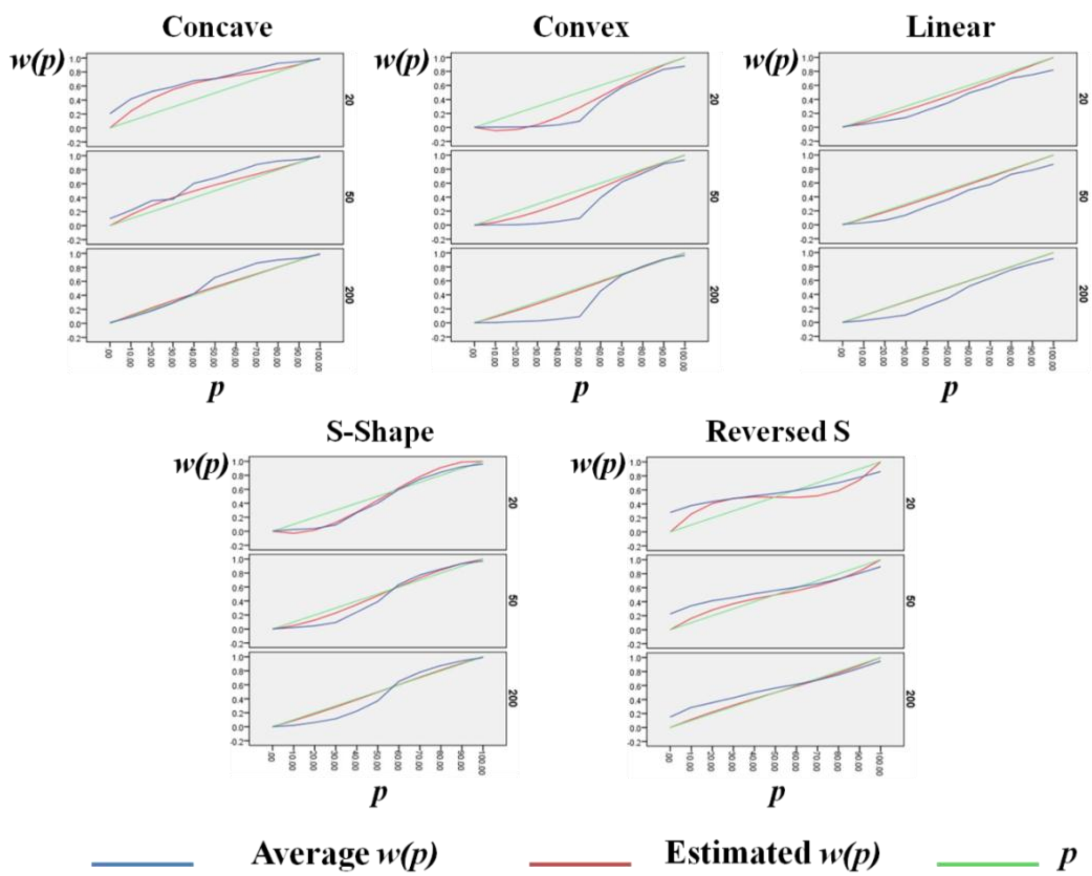


Figure 2.8 Plot of Weighting Functions at Group Levels

The impact of seller review valence and volume. R-square at the group level ranges from .206 for the linear weighting function group to .505 for the *S*-shaped weighting function group above the cross-over point. Supporting H_2, the SR valence has a positive impact on WTP (p-value = .000) for all groups.

As expected, the impact of SR volume on WTP varies across groups. For the concave group, SR volume has a negative impact on WTP (β_N= –1.005 with significance at .000).For the convex group, SR volume has a positive impact on WTP (β_N= .296 with significance at .059).For the linear group, SR volume has no impact on WTP (β_N= –.194 with significance at .139).For the *S*-shaped group, SR volume has no impact on WTP(β_N= –.027 with significance at .874) below the cross-over point, which is not consistent with the hypothesis, but the impact of SR volume is consistent with the hypothesis above the cross-over point (β_N= –.184 with significance at .015).For the reversely *S*-shaped group, SR volume has a negative impact on WTP below the cross-over point (β_N= –.542 with significance at .000) and a positive impact above the cross-over point(β_N= .107 with significance at .030). In general, all hypotheses are supported excepted for the impact of SR volume on WTP for the *S*-shaped group when the SR valences are below the cross-over point. The detailed results are shown in Table 2.4.

DISCUSSION

The results from the experimental study provide relatively strong support for the hypotheses. First, the data confirm that the impact of seller review volume on WTP not only varies across individuals, as maintained by previous research, but also, for some consumers, within an individual depending on the level of review valence. Second, the impact of review volume on WTP is much more complex than previously proposed, because a consumer can

Table 2.4 The Impact of Online Reviews on Consumers' WTP

	Coefficient	Std. Error	t-Stat	p-value	Hypothesis
1. Concave Group (overweight)					
Log(*N*)	-1.005	0.175	-5.748	0.000	Support
Log(*p'*)	0.834	0.110	7.608	0.000	Support
R-square	0.386				
Student #	10				
2. Convex Group (underweight)					
Log(*N*)	0.296	0.155	1.905	0.059	Marginal Support
Log(*p'*)	2.937	0.358	8.211	0.000	Support
R-square	0.498				
Student #	7				
3. **Linear Group (neither underweight nor overweight)**					
Log(*N*)	-0.194	0.131	-1.481	0.139	Support
Log(*p'*)	1.387	0.148	9.381	0.000	Support
R-square	0.260				
Student #	22				
4. ***S*-Shaped Group**					
Below cross-over point (underweight)					
Log(*N*)	-0.027	0.168	-0.158	0.874	Not Support
Log(*p'*)	1.207	0.181	6.651	0.000	Support
R-square	0.324				
Above cross-over point (overweight)					
Log(*N*)	-0.184	0.076	-2.440	0.015	Support
Log(*p'*)	2.055	0.343	5.998	0.000	Support
R-square	0.285				
Student #	25				
5. Reversely *S*-Shaped Group					
Below cross-over point (overweight)					
Log(*N*)	-0.542	0.081	-6.717	0.000	Support
Log(*p'*)	0.359	0.045	8.044	0.000	Support
R-square	0.405				
Above cross-over point (underweight)					
Log(*N*)	0.107	0.049	2.173	0.030	Support
Log(*p'*)	2.591	0.195	13.276	0.000	Support
R-square	0.505				
Student #	51				
			With dummy variables for individuals		

overweight/underweight small/large probabilities or have a reversed pattern. Hence, the influence of volume on WTP can exhibit different patterns among consumers.

In the next essay, I assess the external validity of the framework. I have collected transactional data for Playstation 2 game consoles sold on eBay.com. I expect that the analysis of the empirical data will be much more difficult. Specifically, some covariate variables may need to be controlled. Therefore, additional information related to transactions will be recorded, such as the feedback score (review valence) for consumers who provide reviews to sellers, the time during the day at which the auction ends, the shipping options and other services provided by the seller, and so on. In contrast to the experimental data, it is also difficult to obtain multiple instances of data at the individual level to empirically assess the shape of the weighting function for each individual. Therefore, I plan to conduct my analyses at the segment level. First, consumers can be classified into the different groups using a finite mixture regression model. Second, the linear regression will be performed for each group just as it was for the experimental data. I demonstrate the technique of separating latent consumer groups with the proposed weighting function model, and present the results of the hypotheses testing for each group.

ESSAY THREE. ONLINE REVIEWS AND CONSUMERS WILLINGNESS TO PAY: AN EMPIRICAL INVESTIGATION

INTRODUCTION

Motivation

Websites like eBay heavily depend on their review systems to build trustworthy marketplaces. However, as discussed in Essay One, we still lack clear evidence concerning how consumers use reviews in their purchase decisions for these markets. Many studies examine ratings that combine review volume and review valence, for example, the “Feedback Score” provided by eBay, but these studies have produced mixed results. To understand the role of reviews in consumers’ decision-making processes, it is very important to look at the influence of review valence and review volume separately, the possible interaction between them (Khare et al., 2011; Park et al., 2012), and consumer heterogeneity related to online reviews (Sun, 2012; Wu and Ayala, 2012)

In Essay Two, I proposed that heterogeneity exists among consumers when using seller review information to determine willingness to pay. As a result, there are different interaction patterns between review valence and review volume. While seller review valence should always positively influence consumers’ willingness to pay, review volume varies among consumers, and the preference towards review volume can be described by a consumer’s weighting function. Combining classic expected utility and prospect theory frameworks, I proposed five shapes of weighting functions: concave, convex, linear, *S*-shaped, and reversely *S*-shaped. In an experimental study, I suggested that the preference towards review volume can be very complex. Not only can consumers have totally opposite preferences towards review volume, for some, their preferences also can change according to review valence. My hypotheses are summarized in Table 3.1.

Table 3.1 Summary of Hypotheses

Group	Weighting Function Shape	Description	H1. Impact of Review Valence *p'* on *WTP*	H2. Impact of Review Volume *N* on *WTP*
1	Concave	Overweight all probabilities	+	−
2	Convex	Underweight all probabilities	+	+
3	Linear	Neither underweight nor overweight probabilities	+	No impact
4	*S*-Shaped	Underweight small probabilities; Overweight large probabilities	+	+ Below cross-over point − Above cross-over point
5	Reversely *S*-Shaped	Overweight small probabilities; Underweight large small probabilities	+	− Below cross-over point + Above cross-over point

There are several important considerations that motivate my empirical study. First, testing my theory in online markets is the common approach for establishing its external validity, and thus the relevance of my proposed theory for managerial implications. Second, online markets differ from a lab setting in many aspects. Consumers have different decision goals and processes; furthermore, seller reviews pose different and more challenging distributions. For example, the majority of sellers have review valences close to 100%. On one hand, studies find that people are biased when they process review information, placing more emphasis on review valence and underweighting review volume (Wolf and Muhanna, 2011). On the other hand, because of the large number of high review valences, it becomes less effective in separating good sellers from bad; hence, its impact on price premium becomes less significant (Bockstedt and Goh, 2011). Third, researchers often observe only a few transactions for a given time window.

This approach requires different statistical techniques for constructing variables and analyzing data than those used in my lab setting, which I will elaborate later in this essay.

To provide insightful managerial implications, it is important to test whether the proposed heterogeneity exists in the real online market, and thereby establish the external validity of the framework. Thus, I will test my hypotheses with online transaction data collected from eBay.com. As discussed in Essay Two, it may be difficult to estimate a consumer's weighting function individually when the data lacks sufficient observations from a single consumer. Also, to accommodate consumer differences and at the same time achieve economic efficiency, marketing strategies and activities are often directed toward segments rather than individuals. Hence, for online transactional data, it is more practical to test hypotheses atthe group level. Consistent with the method used in Wu and Ayala (2012), I will first use a finite mixture regression model to segment consumers based on their weighing functions and then test the hypotheses for each group.

This essay contains the following sections. First, I introduce the method, finite mixture regression models, which allows me to simultaneously classify observations into groups using the weighing function model and estimate the parameters for each group. Second, I describe a simulation study that demonstrates the ability of finite mixture regression models to identify the underlying true weighting functions of different groups. Third, I explain my adoption of this method to test the hypotheses with online transaction data. Last, I discuss the study results and future research.

Method

For decades, marketers have used finite mixture regression models, also known as latent class regression models (DeSarbo and Cron, 1988), to identify different segments of consumers whose preferences for marketing information vary. Finite mixture regression models,under the maximum likelihood framework, use the expectation-maximization (EM) algorithm to segment observations into different groups and provide maximum likelihood estimates for model parameters for each group. Based on Frühwirth-Schnatter (2006) and Leisch (2004), the definitions and principles of finite mixture regression models are explained below.[1]

A random variable Y is sampled from a population comprised of K subgroups (usually called components), but group indicators not recorded. All group densities come from the same parametric distribution family with density *f(θ),* where parameter θ differs across groups. Then the conditional density of Y can be shown as below:

$$h(y|x,\varphi) = \sum_{k=1}^{K} \pi_k\, f(y|x,\theta_k)$$

where $\pi_k \geq 0$, $\sum_{k=1}^{K} \pi_k = 1$, $\varphi = (\pi_1, \dots, \pi_k, \theta'_1, \dots, \theta'_k)$, h is the conditional density of y, x is a vector of independent variables, π_k(also called weight distribution) is the prior probability of component k, θ_k is the component-specific parameter vector for the density function f, and φ is the vector of all parameters.

The posterior probability that an observation belongs to component j is specified below. Data can then be segmented by assigning each observation to the component with the maximum posterior probability.

$$P(j|x,y,\varphi) = \frac{\pi_j f(y|x,\theta_j)}{\sum_k \pi_k\, f(Y|x,\theta_k)}$$

The log-likelihood of a sample of N observations is given by the equation below. Because the posterior probability usually cannot be estimated directly, the EM algorithm is used to obtain maximum likelihood estimates of the parameters.

$$logL = \sum_{n=1}^{N} logh(y_n|x_n, \varphi) = \sum_{n=1}^{N} log\,(\sum_{k=1}^{K} \pi_k f(y_n|x_n, \theta_k))$$

The EM algorithm can be used to compute maximum likelihood estimates for incomplete data for which the group indicator is missing. Each iteration of the EM algorithm involves an expectation step (E-step) followed by a maximization step (M-step) (Dempster et al., 1977).

E-step estimates the posterior probability for each observation:

$$\hat{p}_{nk} = P(k|x_n, y_n, \bar{\varphi})$$

and updates prior probability for each component:

$$\hat{\pi}_k = \frac{1}{N}\sum_{n=1}^{N} \hat{p}_{nk}$$

M-step uses the posterior probabilities of each observation as weights for calculating the maximum likelihood estimate for each component:

$$max_{\theta_k} \sum_{n:z_{nk}=1} logf(y_n|x_n, \theta_k)$$

The iteration is repeated until likelihood improvement falls below a pre-specified value or the iteration reaches a maximum number.

In the next section, I explain how a finite mixture regression model was used to separate subjects from simulated samples. The “Flexmix” package (Leisch, 2004) designed for R software was used to apply the finite mixture regression model.

A SIMULATION STUDY

The purpose of the simulation is to assess the ability of the finite mixture model to separate subjects with different weighting functions. The proposed theoretical model for measuring a subject’s attitude weighting function is shown below:

$$\frac{WTP}{V} = c - a\frac{p'(1-p')}{N} - b\frac{p'^2(1-p')}{N}$$

where WTP is the willingness to pay, *V* is the product value, *p'* is seller review valence, and *N* is seller review volume. This model can be fitted regularly as a polynomial regression. However, only three parameters, *a*, *b*, and *c*, need to be estimated; therefore, I used a linear instead of a polynomial regression, as shown in the equation below.

$$Y = c + aX_a + bX_b + \varepsilon; Y = WTP, X_a = -\frac{Vp'(1-p')}{N}, X_b = -\frac{Vp'^2(1-p')}{N}$$

As in Essay Two, the shape of the weighting function is determined by the estimations of parameters *a* and *b*, as shown in Table 2.3.For groups 4 and 5, the cross-over point is determined by –*a*/*b*.

Simulation Data

In this section, I discuss the data used for the simulation.

Data generation. I used the following steps to generate data for each variable:

Review valence *p'*: a random variable that follows a uniform distribution between 0 and 1

Review number *N*: a random variable that follows a uniform distribution between 1 and 2,000

Product value *V*: $800.00

Error *ε*: a random variable that follows a standard normal distribution

Sample size. In Essay Two, I proposed that there are five types of weighting functions. For this simulation, I generated 500 observations for each group; hence the total sample size of the simulated data is 2500.

Parameters. The parameters for each group are shown in Table 3.2.

Table 3.2 Summary of Simulated Parameters

Group	Shape	Parameters		
		c	***a***	***b***
1	Concave	0	–20	0
2	Convex	0	20	0
3	Linear	0	0	0
4	*S*-Shaped*	0	20	–40
5	Reversely *S*-Shaped*	0	–20	40

*The cross-over point for both *S*-shaped and reversely *S*-shaped groups is 0.5.

Testing Scheme

I created five subsets of the simulated data. Subset 1 contained subjects from group 3; subset 2 contained subjects from groups 2 and 3; subset 3 contained subjects from groups 1, 2, and 3; subset 4 contained subjects from groups 1, 2, 3, and 4; and the last subset, 5, contained all of the subjects in the simulated data. The composition of each subset is shown in Table 3.3.

Table 3.3 Summary of Subsets of Simulated Data

Subset	Groups included in the data	Size
1	Group 3	500
2	Group 3, Group 2	1000
3	Group 3, Group 2, Group 1	1500
4	Group 3, Group 2, Group 1, Group 4	2000
5	Group 3, Group 2, Group 1, Group 4, Group 5	2500

A finite mixture regression model was applied to each subset to estimate the parameters for that subset.

Results

The finite mixture regression model generated multiple models with different numbers of components. Under the maximum likelihood framework, Akaike Information Criteria (AIC)

(Akaike, 1974) can be used to choose the best model. As shown below, AIC accounts for both likelihood and model complexity.

$$AIC = -2logL + 2d$$

where L is the likelihood function of the model and d is the number of parameters in the model.

For each subset, the model with the minimum AIC value was selected. For the first four subsets, the finite mixture regression model successfully identified the number of groups embedded in the data. For subset 5, the finite mixture regression model identified six groups instead of five; the extra group, however, had a very small size of 8 observations. See Table 3.4.

Table 3.4 Summary of Selected Models from Each Subset

Subset	Data	Component	Log Likelihood	d.f.	AIC
1	G3	1	–713.5234	4	1435.047
2	G3, G2	2	–1758.183	9	3534.366
3	G3, G2, G1	3	–3031.585	14	6091.170
4	G3, G2, G1,G4	4	–4106.782	19	8251.564
5	G3, G2, G1,G4, G5	6	–5210.928	24	10469.860

Parameter estimation. For each subset, the parameters estimated for each group are shown in Table 3.5. The finite mixture regression model successfully identified all of the groups, and for each group, the estimates were very close to the true value of the parameters. For subset 5, the finite mixture regression model generated6 components; the extra component, component 4, belonged to group 4. Furthermore, the estimates of component 4 were different from the true parameters of group 4. Again, the extra component only had 8 observations, and this result probably was due to random errors.

Hit ratio. The hit ratio for each subset is shown below. When the data contained only one group, the finite mixture regression model correctly identified the group. As one group at a time was added to the data, the overall hit ratio decreased from 100% to 56.92%.

Table 3.5 Parameter Estimations for Simulated Data

Data	Component	Size	Group		Coefficient	Std. Error	Z Value	P
G3	1	500	G3	X_a	0.075	0.079	0.954	0.340
				X_b	–0.151	0.155	–0.975	0.329
G3, G2	1	474	G3	X_a	0.085	0.076	1.110	0.267
				X_b	–0.168	0.150	–1.121	0.263
	2	526	G2	X_a	19.978	0.068	294.661	0.000
				X_b	–0.020	0.186	–0.110	0.912
G3, G2, G1	1	493	G1	X_a	–19.960	0.077	–259.538	0.000
				X_b	–0.112	0.150	–0.744	0.457
	2	522	G2	X_a	19.978	0.067	296.825	0.000
				X_b	–0.020	0.184	–0.106	0.915
	3	485	G3	X_a	0.079	0.071	1.120	0.263
				X_b	–0.158	0.139	–1.140	0.254
G3, G2, G1, G4	1	483	G4	X_a	19.988	0.066	301.679	0.000
				X_b	–39.987	0.129	–311.025	0.000
	2	568	G2	X_a	19.980	0.066	300.862	0.000
				X_b	–0.027	0.182	–0.148	0.882
	3	454	G3	X_a	0.070	0.072	0.967	0.333
				X_b	–0.150	0.142	–1.055	0.291
	4	495	G1	X_a	–19.964	0.077	–260.096	0.000
				X_b	–0.104	0.150	–0.692	0.489
G3, G2, G1, G4, G5	1	554	G1	X_a	–19.959	0.078	–257.394	0.000
				X_b	–0.115	0.151	–0.759	0.448
	2	378	G5	X_a	–19.928	0.126	–157.718	0.000
				X_b	39.838	0.247	161.320	0.000
	3	245	G3	X_a	0.062	0.084	0.728	0.466
				X_b	–0.134	0.164	–0.816	0.414
	4	8	G4	X_a	0.507	0.092	5.505	0.000
				X_b	–0.547	0.148	–3.691	0.000
	5	656	G2	X_a	19.977	0.068	294.868	0.000
				X_b	–0.015	0.186	–0.080	0.936
	6	659	G4	X_a	19.988	0.064	310.890	0.000
				X_b	–39.988	0.123	–324.673	0.000

However, in comparison with the hit ratio of random assignment of subjects to groups, the advantage of the finite mixture regression model became more salient as the number of groups increased. When the data included all five groups, the hit ratio of the finite mixture regression model was almost three times that of the hit ratio of random assignment.

Table 3.6 Hit Ratios of Selected Models

Data	Group	Component					Hit Ratio
		G1	G2	G3	G 4	G5	
	G1						
	G2						
G3	G3			**500**			100%
	G4						
	G5						
	Overall						100%
		G1	G2	G3	G 4	G5	
	G1						
	G2		**441**	59			88.20%
G3, G2	G3		85	**415**			83.00%
	G4						
	G 5						
	Overall						85.60%
		G1	G2	G3	G 4	G5	
	G1	**416**	15	69			83.20%
	G2	14	**432**	54			86.40%
G3,G2,G1	G3	63	75	**362**			72.40%
	G4						
	G5						
	Overall						80.67%
		G1	G2	G3	G4	G5	
	G1	**388**	16	48	48		77.60%
	G2	8	**406**	29	57		81.20%
G3,G2,G1,G4	G3	46	66	**267**	121		53.40%
	G4	53	80	110	**257**		51.40%
	G5						
	Overall						65.90%
		G1	G2	G3	G 4	G5	
	G1	**371**	24	10	61	34	74.20%
	G2	7	**395**	7	71	20	79.00%
G3,G1,G2,G4,G5	G3	43	63	**154**	154	86	30.80%
	G4	55	84	42	**292**	27	58.40%
	G5	78	90	32	89	**211**	42.20%
	Overall						56.92%

Table 3.7 Comparison of Finite Mixture Regression Model and Random Assignment

Data	Hit Ratio of Finite Mixture Regression Model	Hit Ratio of Random Assignment
G3	100.00%	100.00%
G3, G2	85.60%	50.00%
G3, G2, G1	80.67%	33.33%
G3, G2, G1, G4	65.90%	25.00%
G3, G2, G1, G4, G5	56.92%	20.00%

Discussion

The simulation study shows that the finite mixture regression model was very effective at separating subjects into different groups and identifying the true parameters of each group. As the number of underlying groups increased, the method became even more superior.

At the same time, I acknowledge the challenges of using a finite mixture regression model in this particular case. First, the regression model is complex, as shown by the hit ratio, which dropped dramatically as the model's complexity increased. When a quadratic term was introduced to the model by adding the convex group (G2) to the linear group (G3), the hit ratio dropped about 15%, from 100% to 85.6%. Also, when a cubic term was introduced to the model by adding the *S*-shaped group (G4), again, the hit ratio dropped about 15%, from 80.67% to 65.9%. However, if the model already contained a quadratic or cubic term, adding another term of the same power (G1 and G5) led to much smaller decreases in hit ratio. Second, for my simulation data, I generated review valences based on a uniform distribution; however, as discussed in the previous section, samples drawn from eBay usually have high review valences. Sellers who have low review valences either exit the market or change their IDs and rebuild their review profile (Lin et al., 2006; Abbasi et al., 2008). Such a skewed distribution of review valences limits the ability of the finite mixture regression model to identify the true underlying parameters.

AN EMPIRICAL STUDY

eBay's Review System

eBay's rating system has gone through several changes since it was introduced. At the time the data for this study was collected, eBay's review system worked in the following way: A buyer could submit feedback to a seller after each transaction; the feedback could be positive, neutral, or negative. eBay then provided a statistical summary for each member. The "Feedback Score" equaled the number of positive minus the number of negative reviews. The "Positive Feedback Percentage" was the number of positive reviews divided by the sum of positive and negative reviews a member had received in the last 12 months. Both numbers were displayed by the member's login ID, so when a buyer reviewed the auction, she could see the statistics on the same page as the product information. If she clicked the link to visit the seller's profile page, she could view additional information, including the number of positive, neutral, and negative reviews that the seller had received in the past 1, 6, and 12 months; the ratings of the seller for criteria such as communication and shipping time; and detailed comments left by previous customers along with the product they purchased from this seller.

Data Collection

I collected transaction data for anew PlayStation 2 sold on eBay between September and November in 2009. The PlayStation 2 was sold for $299 dollars, and the offline list price did not change during the period of data collection. For each auction, I collected the description of the product; auction information such as shipping policy, return policy, payment policy, etc.; bidding history; and seller profile. Originally, 678 observations were collected; however, some were removed from the data for various reasons. First, some auctions did not result in sales, which led to invalid transactions. Second, some sellers had reviews that were 100% positive, because the positive percentage was calculated based on reviews left within a year. Relying on reviews submitted within a year to calculate the positive percentage significantly increased the proportion of sellers with 100% positive responses. To reduce this bias, I removed the observation if a seller's positive percentage was 100% but her most recent 200 reviews were not uniformly positive. Third, some sellers had 100% positive reviews, but had never sold an item on eBay before, accumulating all of their positive reviews from previous purchases on eBay. Research has shown that reviews for a seller's purchase behavior do not influence purchase price (Zhang, 2006); hence I removed the observation if a seller had never sold a product on eBay prior to the transaction recorded. As a result, I deleted 157 observations, and the final data set contained 529 observations.

Variables

Willingness to pay. Similar to the approach used by Sun and Liu (2010), the winning bid plus the shipping cost were totaled to measure a buyer's willingness to pay for the product. It is reasonable to consider shipping cost when measuring willingness to pay, because when an eBay consumer wins an auction, the amount paid will include the bid price and the shipping cost charged by the seller. Previous research has shown that consumers will consider shipping cost when they participate in auctions and auctions with higher shipping costs usually result in lower final bidding prices (Bockstedt and Goh, 2011).

Review volume *N*. eBay provided a feedback score for each member, which was the difference between the number of positive and negative reviews, instead of the total number of reviews. As discussed above, the feedback score contains information about the review volume and the review valence, which is insufficient for explaining the relationship between reviews and price premium. To consider review valence and review volume separately, and to avoid confounding the two constructs, I measured review volume by estimating the total number of reviews a seller had. Using a formula based on feedback score and positive review percentage, I calculated the number of reviews as shown below:

$$N = positive\ review\ \# + negative\ review\ \# = \frac{2FS(1-p')}{2p'-1} + FS$$

$$FS = positive\ review\ \# - negative\ review\ \#$$

$$p' = positive\ review\ \#/(positive\ review\ \# + negative\ review\ \#)$$

Review valence p'. Review valence is equivalent to the percentage of positive reviews, which was provided by eBay.

Control variables. Variables that also may influence willingness to pay were included in the model as control variables. Some items were featured, or displayed at the top of search results, and some items had special features, such as a warranty. Specialty items may influence the final price because consumers may perceive them as more valuable or less risky than the regular items. Zhou et al. (2009) found that offering a full warranty for the product significantly increases the auction price, and Bockstedt and Goh (2011) found that featured items are sold at a higher price than non-featured items. Therefore, I included a dummy variable, "Specialty," which indicated whether the auction item was listed as a featured item or had special features: 0 denoted a regular item and 1 denoted a specialty item.

Acceptance of returned products reduces the risk associated with a purchase; hence, consumers may pay less for a product if it's non-returnable. I used a dummy variable, "Return," to indicate the return policy of a seller, with 0 denoting that returns were accepted or that information was not provided, and 1 denoting that the seller did not accept returns.

Suter and Hardesty (2005) found that the number of bidders increases as the starting bid set by the seller increases, and as a result, seller's earnings increase. Kamins et al. (2004), who proposed the opposite influence of the starting bid on final price, found that the number of bidders consistently has a positive influence on final bidding price, fully mediating the relationship between the starting bid and the final price. As previous research has shown that the number of bidders in fluencies the final price, I included a variable, “Bidders,” to account for this effect.

It also has been shown that auctions ending during peak time generally have higher closing prices, and that consumers pay more attention to auctions during its closing period regardless of the length or closing day of the auction (Melnik and Alm, 2002).Based on that research, I used a dummy variable, “Hour,” to indicate the peak period of transactions. A value of0 indicated that the auction ended sometime between 11:00 p.m. and 8:00 a.m. central standard time (CST), and that during this period, there were on average 5.3transactionsper hour. A value of 1 indicated that the auction ended between 8:00 a.m. and 11:00 p.m. CST, and that there were on average 32.1 transactions per hour.

Previous studies also have considered the impact of time on product value (Park et al., 2012).The data were collected throughout the three months, and even though the list price of the product did not change during this period, the perceived value of the product could, especially as the holiday season approached. Similar to the approach taken by Wu and Ayala (2012), I used two dummy variables to account for the monthly fluctuation of the perceived value of the product due to external market conditions. One dummy variable indicated auctions that ended in October, and the other indicated auctions that ended in November.

A summary of the variables and a description of the data are shown in Tables 3.8 and 3.9, respectively.

Table 3.8 Summary of Empirical Data Variables

Variable	Measure
WTP	Final Bid plus Shipping Fee charged by the seller
N	Review Volume
p'	Review Valence
Specialty	Whether the item was listed as a featured item on eBay: 0 means no and 1 means yes
Return	Seller's return policy:0 means either accepts returns or does not provide information about return policy and 1 means does not accept return
Bidders	The number of bidders who bid in the auction
Hour	0: low transaction period from 23:00 to 8:59 CST 1: high transaction period from 9:00 to 22:59 CST
Month10	0: auction did not end in October; 1: auction ended in October
Month11	0: auction did not end in November; 1: auction ended in November

Table 3.9 Empirical Data Description

Variable	Mean	Std. Deviation
WTP	302.74	19.16
N	825.45	1551.6
p'	0.9916	0.0163
Bidders	10	4.45
	Number of 0	**Number of 1**
Specialty	482	47
Return	306	223
Hour	48	481
Month10	343	186
Month11	383	146

Analyses

I used a finite mixture regression model to segment 529 observations into different groups,and linear regression models for the observations in each group to test the hypothesis with respect to that group. The models for classifying observations and testing hypotheses are shown below.

Classification model. To classify observations, I used the model proposed in Essay Two, with the addition of the control variables.

$$Y_i = c_i + \beta_a X_{ai} + \beta_b X_{bi} + \beta_1 Specialty_i + \beta_2 Return_i + \beta_3 Hour_i + \beta_4 Month_{10i} + \beta_5 Month_{11i} + \beta_6 Bidders_i + \varepsilon_i$$

where $Y_i = WTP_i - 299 \times p'_i$, $X_{ai} = 299 \times (-p'_i) \times (1 - p'_i)/N_i$,

$X_{bi} = 299 \times (-p'_i) \times p'_i \times (1 - p'_i)/N_i$, and *i*: i^{th}observation

Hypothesis testing model. To test the hypothesis, I used a linear-log function of reviews plus the control variables.[2]As in Essay Two, a linear-log function, was used to instead of a linear function, can capture this diminished return of reputation.

$$LogWTP_{ijk} = c_{ijk} + \beta_N LogN_{ijk} + \beta_p Logp'_{ijk} + \beta_1 Specialty_{ijk} + \beta_2 Return_{ijk} + \beta_3 Hour_{ijk} + \beta_4 Month_{10ijk} + \beta_5 Month_{11ijk} + \beta_6 Bidders_{ijk} + \varepsilon_{ijk}$$

i: i^{th}observation, *j*: j^{th} group, *k*: below or above cross-over point

Aggregate Analysis Results

I ran the hypothesis model with all 529 observations, assuming that there is no difference among consumers in terms of preference towards review volume. Table 3.10 presents the results of the analysis at the aggregate level.

Table 3.10 Aggregate Analysis Results

Variable	Coefficient	Std. Error	t value	P value
Review Valence *p'*	0.728	0.146	4.979	0.000
Review Volume *N*	0.005	0.001	3.623	0.000

At the aggregate level, both review valence and review volume had significant positive impacts on consumers' willingness to pay.

Classification Results

To identify the optimal model, I initially set the pre-specified number of componentsto1, and then increased it to 10one setting at a time. The largest number of components the finite mixture regression model identified was8. The best component models are shown in Table 3.11. I selected the model with the smallest AIC; hence, the 7-component model was selected based on the classification model.

Table 3.11 Model Selection for Empirical Data

Model	# of Components	Log likelihood	d.f.	AIC
1	1	–2255.931	10	4531.863
2	2	–2201.223	21	4444.446
3	3	–2173.237	32	4410.475
4	4	–2152.760	43	4391.519
5	5	–2137.532	54	4383.064
6	6	–2115.566	65	4361.132
7	**7**	**–2098.094**	**76**	**4348.188**
8	8	–2089.303	87	4352.607

The 7-component model identified 3 out of 5 groups: 20.6% of the consumers belonged to the linear group, 38.2% were *S*-shaped, and 41.2% were reversely *S*-shaped. Consistent with the literature and experimental study in Essay Two, the reversely *S*-shaped group was the largest. For the *S*-shaped group, all observations were located above the cross-over point, so within the range of the sample, the *S*-shaped group can be considered a convex group. Detailed information for the 7-component model is shown in Table 3.12.

Hypothesis Testing Results

Below I discuss the impact of review valence and the impact of review volume separately.

Table 3.12 7-Component Model Parameter Estimations

Component	Size	Group		Coefficient	Std. Error	Z Value	P	Cross-over Point
1	54	G3	X_a	6225.269	5053.500	1.232	0.218	NA
			X_b	–6450.339	5293.242	–1.219	0.223	
2	55	G3	X_a	–610.529	2268.808	–0.269	0.788	NA
			X_b	661.048	2397.805	0.276	0.783	
3	112	G5	X_a	–1296.158	275.667	–4.702	0.000	0.9561
			X_b	1355.640	293.296	4.622	0.000	
4	59	G5	X_a	–2685.200	93.154	–28.826	0.000	0.9411
			X_b	2853.300	99.621	28.641	0.000	
5	70	G4	X_a	636.825	201.779	3.156	0.002	0.8912
			X_b	–714.567	214.832	–3.326	0.001	
6	132	G4	X_a	1086.973	305.676	3.556	0.000	0.8830
			X_b	–1231.051	327.791	–3.756	0.000	
7	47	G5	X_a	–1180.900	95.125	–12.414	0.000	0.9375
			X_b	1259.600	100.720	12.506	0.000	

The impact of review valence *p'*. The results showed that, in general, review valence *p'* had a significant positive impact on consumers' willingness to pay. However, in a result inconsistent with the hypothesis, review valence had no impact on willingness to pay for the linear weighting group, and had a negative influence for the reversely *S*-shaped weighting group when it was below the cross-over point. For the rest of the consumers, as held by the hypotheses, review valence showed a positive influence on willingness to pay. With respect to the linear shaped weighting group, out of 109 observations, 53 had a 100% review valence. Therefore, even though the impact of review valence was insignificant, the positive coefficient was still a strong sign of its positive impact on willingness to pay.

The impact of review volume *N*. As expected, the impact of review volume on willingness to pay varied among groups. Consistent with the hypotheses, review volume had no impact on willingness to pay for the linear shaped weighting group. For the *S*-shaped weighting group, review volume showed a negative influence on willingness to pay when review valence was above the cross-over point, although such an effect was statistically insignificant. For the reversely *S*-shaped weighting group, review volume had a negative impact on willingness to pay when the valence was below the cross-over point, but a positive impact when it was above. Table 3.13 presents the results of the hypothesis testing.

Table 3.13 Hypothesis Testing Result Summary

Variable	Group	Coefficient	Std. Error	t Value	P	Hypothesis
Review Valence *p'*	G3	0.156	0.410	0.380	0.705	RD*
	G4 Above Cross-over Point	0.244	0.120	2.031	0.044	S
	G5 Below Cross-over Point	–1.631	0.447	–3.652	0.022	NS
	G5 Above Cross-over Point	2.136	0.197	10.856	0.000	S
Review Volume *N*	G3	0.004	0.005	0.821	0.414	S
	G4 Above Cross-over Point	–0.001	0.001	–1.015	0.311	RD
	G5 Below Cross-over Point	–0.039	0.008	–5.002	0.007	S
	G5 Above Cross-over Point	0.006	0.001	4.364	0.000	S

* RD: Estimate had same sign as proposed by hypothesis, but effect was not significant.
S: Hypothesis was supported at significant level of 0.05.
NS: Hypothesis was not supported at significant level of 0.05.

DISCUSSION

Both the experimental and empirical studies confirmed that consumer heterogeneity exists and influences the way that consumers use seller review information in their purchase decisions. Although the empirical study only identified3 out of the 5 groups originally proposed, it showed that consumers can be very different in their preferences towards review volume: some consumers simply do not care much about review volume, some consumers have relatively stable preferences towards review volume, and some consumers will change their preferences towards review volume based on review valence. On the aggregate level, my empirical data showed that review volume has a positive effect on consumers' willingness to pay, because the majority of the observations fell in the reversely *S*-shaped group and review valence was above the cross-over point. I expect that the relationship between review volume and consumers' willingness to pay will change if the sample's composition changes. Therefore, it is reasonable to expect inconsistent observations of the influence of review volume on an aggregate level when the conclusions are drawn from different samples.

The limitation of the current empirical study is that the data collection period was not long enough to identify sellers with low review valences, because these sellers may eventually be eliminated by the market. As a result, the distribution of review valence was negatively skewed, and it is hard to identify observations below the cross-over points for the *S*-shaped and reversely *S*-shaped groups. Future research can improve the validity of the framework by adopting a larger and more representative set of data.

My research provides a descriptive framework that shows that consumers have different preferences towards review volume and, furthermore, that such differences can be categorized by consumers' weighting functions. My studies establish correlation rather than a causal relationship between weighting function and the impact of review volume on willingness to pay. Future studies can establish a causal relationship by developing independent measurements of weighting functions.

Finally, the current framework was developed under the binary review format; thus it only considers review volume and valence. Future research can extend the framework to include a continuous review format, such as Amazon.com's, and incorporate the influence of review variance on consumers' willingness to pay.

Notes

1. Formulas for finite mixture models are consistent with those shown in Leish (2004).

2. For the observations in group 5 that were below the cross-over point, covariant variables were excluded due to the small sample size.

REFERENCES

Abbasi, Ahmed, Hsinchun Chen and Jay Nunamaker Jr. (2008), "Stylometric Identification in Electronic Markets: Scalability and Robustness," *Journal of Management Information Systems*, 25, 49-78.

Aiken, Damon and David Boush (2006), "Trustmarks, Objective-Source Ratings, and Implied Investments in Advertising: Investigating Online Trust and the Context-Specific Nature of Internet Signals," *Journal of Academy of Marketing Science*, 34, 308-323.

Akaike, Hirotugu (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 12, 716-723.

Aperjis, Christina and Ramesh Johari (2010), "Optimal Windows for Aggregating Ratings in Electronic Marketplaces," *Management Science*, 56, 864-880.

Awad, Neveen and Arik Ragowsky (2008), "Establishing Trust in Electronic Commerce through Online Word of Mouth: An Examination across Genders," *Journal of Management Information Systems*, 24, 101-121.

Ba, Sulin and Paul Pavlou (2002), "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly*, 26, 243-268.

Bakos, Yannis and Chrystanthos Dellarocas (2011), "Cooperation without Enforcement? A Comparative Analysis of Litigation and Online Reputation as Quality Assurance Mechanisms," *Management Science*, 51, 1944-1962.

Basuroy, Suman, Subimal Chatterjee, and Abraham Ravid (2003), "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," *Journal of Marketing*, 67, 103-117.

Bockstedt, Jesse and Kim Huat Goh (2011), "Seller Strategies for Differentiation in Highly Competitive Online auction Markets," *Journal of Management Information Systems*, 28, 235-267.

Bolton, Gary, Claudia Loebbecke, and Axel Ockenfels (2008), "Does Competition Promote Trust and Trustworthiness in Online Trading? An Experimental Study," *Journal of Management Information Systems*, 25, 145-169.

Bruce, Norris, Ernan Haruvy, and Ram Rao (2004), "Seller Rating, Price, and Default in Online Auctions," *Journal of Interactive Marketing*, 18, 37-50.

Camerer, Colin and Martin Weber (1992), "Recent Developments in Modeling Preference: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 5, 325-370.

Chan, Tat, Vrinda Kadiyali, and Young-Hoon Park (2007), "Willingness to Pay and Competition in Online Auctions," *Journal of Marketing Research*, 44, 324-333.

Chen, Yubo, Scott Fay, and Qi Wang (2011), "The Role of Marketing in Social Media: How Online Consumer Reviews Evolve," *Journal of Interactive Marketing*, 25, 85-94.

Chen, Yubo, Yong Li, and Jurui Zhang (2012), "When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do? The Case of Media Critics and Professional Movie Reviews," *Journal of Marketing*, 75, 116-134.

Chen, Yubo, Qi Wang, and Jinhong Xie (2011), "Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning," *Journal of Marketing Research*, 48, 238-254.

Chen, Yubo and Jinhong Xie (2008), "Online Consumer Review: Word-of-Mouth as A New Element of Marketing Communication Mix," *Management Science*, 54, 477-491.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43, 345-354.

Chevalier, Judith and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43, 345-354.

Chintagunta, Pradeep, Shyam Gopinath, and Sriram Venkataraman (2010), "The Effects of Online User Reviews on Movei Box Office Performance: Accounting for Sequential Rollout and Aggregation across Local Markets," *Marketing Science*, 29, 944-957.

Clemons, Eric, Guodong Gao, and Lorin Hitt (2006), "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems*, 23, 149-171.

De Maeyer, Peter and Hooman Estelami (2011), "Consumer Perceptions of Third Party Product Quality Ratings," *Journal of Business Research*, 64, 1067-1073.

Decker, Reinhold and Michael Trusov (2010), "Estimating Aggregate Consumer Preferences from Online Product Reviews," *International Journal of Research in Marketing*, 27, 293-307.

Dellarocas, Chrysanthos (2005), "Reputation Mechanism Design in Online Trading Environemnts with Pure Moral Hazard," *Information Systems Research*, 16, 209-230.

Dellarocas, Chrysanthos, Guodong Gao, and Ritu Narayan (2010), "Are Consumers More Likely to Contribute Online Reviews for Hit or Niche Products?" *Journal of Management Information Systems*, 27, 127-157.

Dellarocas, Chrysanthos, Xiaoquan Zhang, and Neveen F. Awad (2007), "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing*, 21, 23-45.

Dempster, Arthur, Nan Laird, and Donald Rubin (1977), “Maximum Likelihood from Incomplete Data via the EM-Algorithm,” *Journal of Royal Statistical Society,* 39, 1-38.

DeSarbo, Wayne and William Corn (1988), “A Maximum likelihood methodology for clusterwise linear regression,” *Journal of Classification*, 5, 249-282.

Dewally, Michaël and Louis Ederington (2006), “Reputation, Certification, Warranties, and Information as Remedies for Seller-Buyer Information Asymmetries: Lessons from the Online Comic Book Market,” *Journal of Business,* 79, 693-729.

Dewan, Sanjeev and Vernon Hsu (2004), “Adverse Selection in Electronic Markets: Evidence from Online Stamp Auctions,” *The Journal of Industrial Economics*, 52, 497-516.

Duan, Wenjing, Bin Gu, and Andrew Whinston (2008), “The Dynamics of Online Word-of-Mouth and Product Sales-An Empirical Investigation of the Movie Industry,” *Journal of Retailing*, 84, 233-242.

Duan, Wenjing, Bin Gu, and Andrew Winston (2009), “Informational Cascades and Software Adoption on The Internet: An Empirical Investigation,” *MIS Quarterly,* 33, 23-48.

Eaton, David H. (2005), “Valuing Information: Evidence from Guitar Auctions on eBay,” *Journal of Applied Economics and Policy,* 24, 1-19.

Ellsberg, Daniel (1961), “Risk, Ambiguity and the Savage Axioms,” *Quarterly Journal of Economics*, 49, 643-669.

Einhorn, Hillel and Robin Hogarth (1985), “Ambiguity and Uncertainty in Probabilistic Inference,” *Psychological Review*, 92, 433-461.

Einhorn, Hillel (1986), “Decision Marking under Ambiguity,” *Journal of Business*, 59, 225-250.

Feng, Jie and Purushottam Papatla (2011), “Advertising: Stimulant or Suppressant of Online Word of Mouth?” *Journal of Interactive Marketing*, 25, 75-84.

Finch, Byron (2007), “Customer Expectations in Online Auction Environments: An Exploratory Study of Customer Feedback and Risk,” *Journal of Operations Management,* 25. 985-997.

Frühwirth-Schatter, Sylvia (2006), *Finite Mixture and Markov Switching Models*. Springer-Verlag Series in Statistics. Spring-Verlag, New York.

Ghose, Anindya (2009), “Internet Exchanges for Used Goods: An Empirical Analysis of Trade Patterns and Adverse Election,” *MIS Quarterly*, 33, 263-291.

Gilkeson, James and Kristy Reynolds (2003), “Determinants of Internet Auction Success and Closing Price: An Exploratory Study,” *Psychology & Marketing*, 20, 537-566.

Godes, David and Dina Mayzlin (2004), “Using Online Conversations to Study Word of Mouth Communication,” *Marketing Science*, 23, 545-560.

Gonzalez, Richard and George Wu (1999), “On the Shape of Weighting Function,” *Cognitive Psychology*, 38, 129-166.

Hastie, Reid and Robyn M. Dawes (2001), *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making*, 2nd Edition, SAGE Publications, California.

Henning-Thurau, Thorsten, Kevin Gwinner, Gianfranco Walsh, and Dwayne Gremler (2009), “Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on The Internet?” *Journal of Interactive Marketing*, 18, 38-52.

Houser, Daniel and John Wooders (2006), “Reputation in Auctions: Theory, and Evidence from eBay,” *Journal of Economics and Management Strategy*, 15, 353-369.

Hogarth, Robin and Hillel Einhorn (1990), “Venture Theory: A Model of Decision Weights,” *Management Science*, 36, 780-803.

Hogarth, Robin and Howard Kunreuther (1992), “Pricing Insurance and Warranties: Ambiguity and Correlated Risks,” *The GENEVA Papers on Risk and Insurance Theory*, 17, 35-60.

Huang, Yingtang, Jongrong Chen and Wencheng Liu (2011), “Effect of Seller Reputation on Success at Online Auctions: An Empirical Study,” *International Journal of Management*, 28, 621-629.

Huang, Peng, Nicholas Lurie, and Sabyasachi Mitra (2009), “Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods,” *Journal of Marketing*, 73, 55-69.

Kahn, Barbara E. and Rakesh K. Sarin (1988), “Modeling Ambiguity in Decisions under Uncertainty,” *Journal of Consumer Research*, 15, 265-272.

Kahneman, Daniel and Amos Tversky (1979), “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 47, 263-291.

Khare, Adwait, Lauren Labrecque, and Anthony Asare (2011), “The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings,” *Journal of Retailing*, 87, 111-126.

Kim, Junyong and Pranjal Gupta (2011), “Emotional Expressions in Online User reviews: How They Influence Consumers’ Product Evaluations,” *Journal of Business Research*, 65, 985-992.

Kalyanam, Kirthi and Shelby McIntyre (2001), “Return on Reputation in Online Auction Markets,” *Working Paper*, Santa Clara University.

Koh, Noi Sian, Nan Hu, and Eric Clemons (2010), “Do Online Reviews Reflect a Product’s True Perceived Quality? An Investigation of Online Movie Reviews across Cultures,” *Electronic Commerce Research and Applications*, 9, 374-385.

Korfiatis, Nikolaos, Elena Garcia-Bariocanal, and Salvador Sanchez-Alonso (2011), “Evaluating Content Quality and Helpfulness of Online Product Reviews: The Interplay of Review Helpfulness Vs. Review Content,” *Electronic Commerce Research and Applications*, 11, 205-217.

Kumar, Nanda and Izak Benbasat (2006), “Influence of Recommendations and Consumer Reviews on Evaluations of Websites,” *Information Systems Research*, 17, 425-439.

Lee, Jumin, DoHyung Park, and Ingoo Han (2008), “The Effect of Negative Online Consumer Reviews on Product Attitude: An Information Processing View,” *Electronic Commerce Research and Applications*, 7, 341-352.”

Lee, Jung and Jae-Nam Lee (2009), “Understanding the Product Information Inference Process in Electronic Word-of-Mouth: An Objectivity-Subjectivity Dichotomy Perspective,” *Information & Management*, 46, 302-311.

Lee, Mira and Seounmi Youn (2009), “Electronic Word of Mouth (eWOM)-How eWOM Platforms Influence Consumer Product Judgment,” *International Journal of Advertising*, 28, 473-499.

Lee, Thomas and Eric Bradlow (2011), “Automated Marketing Research Using Online Customer Reviews,” *Journal of Marketing Research*, 48, 881-894.

Lei, Qin (2011), “Financial Value of Reputation: Evidence from The eBay Auctions of Gmail Invitations,” *The Journal of Industrial Economics*, 59, 422-456.

Leish, Friedrich (2004), “FlexMix; A General Framework for Finite mixture Models and Latent Class Regression in R,” *Journal of Statistical Software*, 11, 1-18.

Li, Jin and Lingjing (2011), “Online Persuasion: How the Written Word Drives WOM Evidence from Consumer-Generated Product Reviews,” *Journal of Advertising Research*, 51, 239-257.

Li, Shibo, Kannan Srinivasan, and Baohong Sun (2009), “Internet Auction Features as Quality Signals,” *Journal of Marketing*, 73, 75-92.

Li, Xinxin and Lorin Hitt (2008), “Self-Selection and Information Role of Online Product Reviews,” *Information Systems Research*, 19, 456-474.

Li, Xinxin and Lorin Hitt (2010), “Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis,” *MIS Quarterly*, 34, 809-832.

Lin, Zhangxi, Dahui Li, Balaji Janamanchi and Wayne Huang (2006), "Reputation Distribution and Consumer-to-Consumer Online Auction Market Structure: An Exploratory Study," *Decision Support System*, 41, 435-448.

List,John (2004), "Neoclassical Theory Versus Prospect Theory: Evidence from The Marketplace," *Econometrica*, 72, 615-625.

Livingston, Jeffery (2005), "How Valuable Is A Good Reputation? A Sample Selection Model of Internet Auctions," *The Review of Economics and Statistics*, 87, 453-465.

Liu, Yong (2006), "Word of Mouth for Moveis: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70, 74-89.

Lucking-Reiley, David, Doug Bryan, Naghi Prasad, and Daniel Reeves (2007), "Pennies from eBay: The Determinants of Price in Online Auctions," *The Journal of Industrial Economics*, 55, 223-233.

Melnik, Mikhail and James Alm (2002), "Does A Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions," *Journal of Industrial Economics,* 3, 337-349.

Moe, Wendy and Michael Trusov, 2011. "The Value of Social Dynamics in Onlien Product Ratings Forums," *Journal of Marketing Research*, 48, 444-456.

Moon, Sangkil, Paul Bergey, and Dawn Iacobucci (2010), "Dynamic Effects Among Movie Ratings, Movie Revenues, and Viewer Satisfaction," *Journal of Marketing*, 74, 108-121.

Mudambi, Susan and David Schuff (2010), "What Makes A Helpful Online Review? A Study of Customer Reviews on Amazon.com," *MIS Quarterly*, 34, 185-200.

Nayor, Rebecca, Cait Lamberton, and David Norton (2011), "Seeing Ourselves in Others: Reviewers Ambiguity, Egocentric Anchoring, and Persuasion," *Journal of Market Research*, 48, 617-631.

Nielsen (2010), *Global Trends in Online Shopping,* www.hk.hielsen.com/documents/Q12010OnlineShoppingTrendsReport.pdf.

Nielsen (2012), *Global Trust in Advertising and Brand Messages*, www.fi.nielsen.com/site/documents/NielsenTrustinAdvertisingGlobalReportApril2012.pdf.

Obloj, Tomasz and Laurence Capron (2011). "Role of Resource Gap and Value Appropriation: Effect of Reputation Gap on Price Premium in Online Auctions," *Strategic Management Journal*, 32, 447-456.

Pan, LeeYun and JyhShen Chiou (2011), "How Much Can You Trust Online Information? Cues for Perceived Trustworthiness of Consumer-Generated Online Information," *Journal of Interactive Marketing*, 25, 67-74.

Park, Do-Hyung and Sara Kim (2008), "The Effects of Consumer Knowledge on Message Processing of Electronic Word-of-Mouth via Online Consumer Reviews," *Electronic Commerce Research and Applications*, 7, 399-410.

Park, JaeHong, Bin Gu, and HoonYoung Lee (2011), "The relationship between Retailer-Hosted and Third-Party Hosted WOM Sources and Their Influence on Retailer Sales," *Electronic Commerce Research and Applications*, 10, 304-312.

Park, Yong-Hoon and Eric Bradlow (2005), "An Integrated Model for Bidding Behavior in Internet Auctions: Whether, Who, When, and How Much," *Journal of Marketing Research*, 42, 470-482.

Pavlou, Oaul and Angelika Dimoka (2006), "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation," *Information Systems Research*, 17, 392-414.

Prelec, Drazen (1998), "The Probability Weighting Function," *Econometrica*, 66, 497-527.

Sarin, Rakesh K. and Martin Weber (1993), "Effects of Ambiguity in Market Experiments," *Management Science*, 39, 602-615.

Schlosser, Ann (2011), "Can Including Pros and Cons Increase the Helpfulness and Persuasiveness of Online Reviews? The Interactive Effects of Ratings and Arguments," *Journal of Consumer Psychology*, 21, 226-239.

Schmidt, Ulrich, Chris Starmer, and Robert Sugden (2008), "Third-Generation Prospect Theory," *Journal of Risk Uncertainty*, 36, 203-223.

Sun, Monic (2012), "How Does the Variance of Product Ratings Matter?" *Management Science*, 58, 696-707.

Standifird, Stephen (2001), "Reputation and E-Commerce: eBay Auctions and the Asymmetrical Impact of Positive and Negative Ratins," *Journal of Management*, 27, 279-295.

Sun, Chia-Hung and Kang Liu (2010), "Information Asymmetry and Small Business in Online Auction Market," *Small Business Economics*, 34, 433-444.

Suter, Tracy and David Hardesty (2005), "Maximizing Earnings and Price Fairness Perceptions in Online Consumer-to-Consumer Auctions," *Journal of Retailing*, 81, 307-317.

Taylor, David and David Strutton (2010), "Has E-Marketing Come of Age? Modeling Historical Influences on Post-Adoption Era Internet Consumer Behaviors," *Journal of Business Research*, 63, 950-956.

Tirunillai, Seshadri and Gerard Tellis (2012), "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science*, 31, 198-215.

Tversky, Amos and Daniel Kahneman (1986), "Rational Choice and the Framing of Decisions," *Journal of Business,* 59, 251-278.

Tversky, Amos and Daniel Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty,* 5, 297-323.

Tversky, Amos and Craig Fox (1995), "Weighting Risk and Uncertainty," *Psychological Review*, 102, 269-283.

Tversky, Amos and Craig Fox (1998), "A Belief-Based Account of Decision under Uncertainty," *Management Science,* 44, 879-895.

Tversky, Amos and Peter Wakker (1995), "Risk Attitudes and Decision Weights," *Econometrica,* 63, 1255-1280.

Utz, Sonja, Peter Kerkhof, and Joost van den Bos (2012), "Consumers Rule: How Consumer Reviews Influence Perceived Trustworthiness of Online Stores," Electronic Commerce Research and Applications, 11, 49-58.

Weber, Martin and Colin Camerer (1987), "Recent Developments in Modeling Preferences under Risk," *OR Spektrum*, 9, 129-151.

Wolf, James and Waleed Muhanna (2011), "Feedback Mechanisms, Judgment Bias, and Trust Formation in Online Auctions," *Decision Sciences*, 42, 43-68.

Wu, George and Richard Gonzalez (1996), "Curvature of the Probability Weighting Function," *Management Science*, 42, 1676-1690.

Wu, Jianan and Edgardo Arturo Ayala Gaytάn (2012), "The Role of Online Seller Reviews and Product Price on Buyers' Willingness-to-Pay: A Risk Perspective," *European Journal of Information Systems* Advance Online Publication 3 July, doi: 10.1057/ejis.2012.33.

Wu, Jianan, Jie Sun, Yinglu Wu, and Zhilin Yang (2012), "How Much Are Buyers Willing to Pay? A Two Stage Model of Uncertainty Assessment with Online User Reviews," *working paper,* Louisiana State University.

Yao, Song and Carl Mela (2008), "Online Auction Demand," Marketing Science, 27, 861-885.

Yang, Jian, Xiaorui Hu, and Han Zhang (2007), "Effects of a Reputation Feedback System on An Online Consumer-to-Consumer Auction Market," *Decision Support System,* 44, 93-1045.

Yang, Jun and Enping Mai (2010), "Experiential Goods with Network Externalities Effects: An Empirical Study of Online Rating System," *Journal of Business Research,* 63, 1050-1057.

Zeithammer, Robert (2006), “Forward-Looking Bidding in Online Auctions,” *Journal of Marketing Research*, 43, 462-476.

Zhang, Jie (2006), “The Roles of Players and Reputation: Evidence from eBay Online Auctions,” *Decision Support System*, 42, 1800-1818.

Zhou, Ming, Martin Dresner, and Robert Windle (2009), “Revisiting Feedback Systems: Trust Building in Digital Markets,” *Information & Management*, 46, 279-284.

Zhu, Feng and Xiaoquan Zhang (2010), “Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics,” *Journal of Marketing*, 74, 133-148.

APPENDIX: LIST OF LITERATURE REVIEW JOURNALS

Academy of Management Journal

Academy of Management Review

Administrative Science Quarterly

American Economic Review

Decision Sciences

Econometrica

Electronic Commerce Research and Applications

Harvard Business Review

Industrial Marketing Management

Information and Management

Information System Research

International Business Review

International Journal of Advertising

International Journal of Management Review

International Journal of Marketing Research

International Journal of Research in Marketing

International Marketing Review

Journal of Advertising Research

Journal of Business Economics and Management

Journal of Business and Industrial Marketing

Journal of Business and Psychology

Journal of Business Research

Journal of Consumer Psychology

Journal of Consumer Research

Journal of Economic Literature

Journal of Finance

Journal of Information Technology

Journal of Informetrics

Journal of International Business Studies

Journal of International Marketing

Journal of Interactive Marketing

Journal of Management

Journal of Management Information Systems

Journal of Management Studies

Journal of Marketing

Journal of Marketing Research

Journal of Operations Management

Journal of Retailing

Journal of Service Research

Journal of the Academy of Marketing Science

Management Science

Marketing Letter

Marketing Science

MIS Quarterly

MIT Sloan Management Review

Omega-The International Journal of Management Science

Organizational Research Methods

Psychology and Marketing

Research in Organizational Behavior

Review of Economics and Statistics

Review of Financial Studies

Strategic Management Journal

Technological and Economic Development of Economy

The Academy of Management Annals

The Journal of Economic Perspectives

The Quarterly Journal of Economics

The Review of Economic Studies

VITA

Yinglu (Elle) Wu received her bachelor's degree in business administration from Hubei University in China. In 2006, Yinglu came to Louisiana State University to pursue her Marketing PhD, joining the Master of Applied Statistics program in 2010. Yinglu earned her Master's in Applied Statistics in May 2012 and her Doctor of Philosophy degree in Business Administration (Marketing) in December of the same year. In 2012, Yinglu began her career as an assistant professor of marketing at University of Wisconsin-Stevens Point. Her research interests include consumer behavior in online markets, e-commerce, and empirical marketing research. Her research has been presented at conferences such as INFORMS Marketing Science Conference. Yinglu is also actively engaged in academic activities: she was an American Marketing Association Sheth Doctoral Consortium Fellow, a Society of Marketing Advances Doctoral Consortium Fellow, a contributor to an entrepreneur and business book, and a reviewer for a marketing textbook.