Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing 33 (2016) 43 – 56

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

The Impact of User Review Volume on Consumers' Willingness-to-Pay: A Consumer Uncertainty Perspective

Yinglu Wu [a] & Jianan Wu [b,*]

[a] Boler School of Business, John Carroll University, 1 John Carroll Blvd, University Heights, OH 44118, USA
[b] E. J. Ourso College of Business, Louisiana State University, 2113 Business Education Complex, Nicholson Extension, Baton Rouge, LA 70803, USA

Available online 3 February 2016

Abstract

Numerous studies have investigated the impact of review volume and review variance on product price, but their findings are mixed. The perspective of mismatch cost framework (e.g., Sun 2012) argues that, at the market level, the impact of review variance on product price varies with review valence due to diverse product tastes *across* individuals. The perspective of classic expected utility framework (e.g., Wu and Ayala Gaytán 2013; Wu et al. 2013) further argues that heterogeneous risk attitudes *across* individuals directly drives the varying impact of both review volume and variance on willingness-to-pay, regardless of review valence. Although both frameworks have gained good empirical support, neither of them probed whether the impact of review volume or review variance varies *within* an individual.

We extend the current research by focusing on the varying impact of review volume on consumers' willingness-to-pay. Combining economic and behavioral theories of decisions under uncertainty, we argue that consumers' preferences of uncertainty can vary both *across* and *within* individuals. The extended framework thus concludes that the impact of review volume on consumers' willingness-to-pay not only varies across individuals with different types of uncertainty preferences, but may also change with review valence within an individual of some types of uncertainty preferences. The framework is tested using an experimental study and an empirical study. Results from both studies provide good support for this broader framework.

Keywords: Online user review; Willingness-to-pay; Prospect theory

Introduction

Online user reviews are important for both online and offline businesses. Consumers trust online reviews more than advertisements (Nielsen 2012) and rely on online reviews when choosing products (Nielsen 2010) and sellers (Anderson 2014). Among all the information that websites provide for review purposes, statistics are the most prominent and often the first that consumers examine. Key statistics include review volume (the number of reviews a product or a seller receives), review valence (the average review rating), and review variance (usually shown through the distribution of reviews). These statistics serve as a tool for consumers to reduce the burden of information search and decision making (Jang, Prasad, and Ratchford 2012), and hence significantly impact a company's marketing performance. High review volume can increase the exposure of a business or product offering. For example, Yelp uses review statistics to rank businesses in its search results (Holloway 2011). High review valance also increases product consideration (Jang, Prasad, and Ratchford 2012). On average, Google ads with seller ratings get a 17% higher click-through rate than the same ads without review ratings (Friedman 2011).

To date, marketing researchers have invested great effort in studying the impact of review statistics; however, their findings were inconsistent. For example, while review valence is known to positively affect product price and sales (e.g., Chevalier and Mayzlin 2006; Moe and Trusov 2011; Wu and Ayala Gaytán 2013), review volume and variance can have a positive, insignificant, or even negative influence on marketing outcomes (e.g.,

* Corresponding author.
E-mail addresses: ywu@jcu.edu (Y. Wu), jiananwu@lsu.edu (J. Wu).

http://dx.doi.org/10.1016/j.intmar.2015.11.001
1094-9968/

Basuroy, Chatterjee, and Ravid 2003; Chintagunta, Gopinath, and Venkataraman 2010; Clemons, Gao, and Hitt 2006; Zhu and Zhang 2010). A key reason for such contradictory findings is that most of the studies are based on the assumption that online reviews have equal impact on different consumers (King, Racherla, and Bush 2014). While it is important to design marketing strategies that leverage online reviews (Simonson and Rosen 2014), companies must consider consumer heterogeneity in processing review statistics to fully exploit online reviews. As King, Racherla, and Bush (2014) proposed, future research needs to identify and quantify the disaggregate effects of online word-of-mouth. To answer this research call, we conducted a study to unveil the fundamental role of review statistics in consumers' willingness-to-pay (WTP) decisions; to investigate consumer heterogeneity in processing review statistics; and to quantify the varying consumer preferences toward review statistics.

We organize our paper as follows. First, we provide background on consumers' preferences regarding online reviews and describe our contributions to the current literature. Second, we introduce our conceptual framework and develop the hypotheses. Third, we test the hypotheses using two studies: first, an experimental study, and then an empirical study. Accompanying study results are presented in each section. Last, we discuss the managerial implications and limitations of the current research.

Related Literature

Two streams of research motivate our work: research that examines individual preference regarding review statistics, and research that examines interactions among review statistics at the aggregate level. Based on expected utility theory, Wu and Ayala Gaytán (2013) and Wu et al. (2013) suggest that review volume, review valence, and review variance independently impact a consumer's judgment of purchase risk. At the individual level, each consumer has a consistent preference regarding review volume and review variance, but such preference varies from one consumer to another based on each consumer's risk attitude. Using data from eBay.com and Amazon.com, these authors find that high review volume or low review variance positively affect consumer WTP when consumers are risk-averse, but negatively affect WTP when consumers are risk-seeking; for risk-neutral consumers, neither volume nor variance has an impact on WTP.

These studies explore the theoretical foundation of consumer heterogeneity in processing online reviews, but the framework fails to explain an important observation by studies that report interactions among review statistics at the aggregate level. For example, based on mismatch cost, Sun (2012) proposes a model that assumes consumers are risk-neutral with heterogeneous product tastes. Hence, at the individual level, consumers have the same attitude toward review variance; but at the aggregate level, review valence moderates the influence of review variance due to cross-individual differences in product taste. Sun (2012) finds that on Amazon.com, a larger review variance can increase book sales if (and only if) review valence is small. Some experimental studies also suggest the existence of interactions among review statistics at the aggregate level. For example, Khare, Labrecque, and Asare (2011) report a significant main effect of review valence and an interaction between review valence and volume. The impact of review volume on consumer preference is positive when review valence is high, and negative when valence is low.

In this study, we seek to develop a comprehensive framework that can capture the interactive effects of review valence, volume, and variance at the individual level while accounting for consumers' varying preferences as to review statistics. Our framework differs from previous research in several ways. First, our framework shows that the interactions among valence, volume, and variance can occur not only at the aggregate level—as observed by studies such as Sun (2012) and Khare, Labrecque, and Asare (2011)—but also at the individual level, an effect unaddressed in the literature. Second, Sun's mismatch cost theory assumes that consumers' product tastes are uniformly distributed. This assumption may be appropriate for product preference, but it is overly strong for evaluating the seller. Similar to Wu and Ayala Gaytán (2013) and Wu et al. (2013), our framework is built on theories of decisions under uncertainty, and hence can be applied to both product and seller reviews. Last, our framework allows for, but does not require, consistent risk attitude. Unlike Wu and Ayala Gaytán (2013) and Wu et al. (2013), we investigate inconsistent uncertainty preferences, an important characteristic not yet examined in the literature of online user reviews.

We study how consumers use statistics in online reviews to form their WTP toward different online sellers. Similar to Miller et al. (2011), we view a consumer's WTP as a point measure and define it as the maximum price a consumer is willing to pay an online seller for a product, given the option of buying the same product at a fixed price from an offline retailer with no transaction risk. Because purchases from online sellers are associated with high transaction risks, consumers are willing to compensate sellers in order to reduce transaction risks, according to Ba and Pavlou (2002).

We focus on WTP because the impact of user reviews on price is inconclusive and a fuller understanding of this relationship contains direct implications for enhancing targeted pricing and promotion strategies—an important area of research in online marketing that calls for more attention (Grewal et al. 2010). We attend to seller reviews because consumers in online markets discern more uncertainty about sellers than about products (Wu et al. 2013). While consumers can obtain information on product quality from multiple sources (e.g., product commercials and offline retailers), they must obtain information on online seller quality (e.g., eBay and Amazon sellers) primarily—if not entirely—from user reviews. This singular characteristic of seller reviews helps us isolate the impact of online user reviews.

We test our theoretical framework using two methods: an experimental study with participants, and an empirical study using data from eBay.com. Results reveal that the impact of review statistics (e.g., review volume) on consumer WTP varies across consumers, suggesting that online sellers can incorporate review statistics and consumer characteristics into pricing strategies at segment or individual levels.

Conceptual Framework & Hypothesis Development

Uncertainty and Online Purchase Decision

Without seller uncertainty, we assume that V and $v(V)$ are the perceived benefits of a product that a consumer purchases from an offline seller and from an online seller respectively. With online seller uncertainty, we represent the consumer's online purchase by a prospect $(V, p; 0, 1 - p)$, adopting the prospect theory framework (e.g., Tversky and Kahneman 1992): the consumer gains $v(V)$ if she is satisfied with the purchase and 0 if she is dissatisfied, where p is the true probability of reaching a satisfied outcome. Thus, the consumer's WTP to the online seller is her valuation of the prospect $(V, p; 0, 1 - p)$.

In general, a consumer does not know the true outcome probability p, but she can estimate it using the seller reviews she observes, which represent a sample of the review population. Review valence p' is the sample probability and review volume N is the sample size. eBay's review format, the bidirectional user review, is one of the most popular formats for online seller reviews. In a bidirectional review system, we assume that a consumer leaves a positive review if she is satisfied and a negative review if she is not. If a seller has N bidirectional reviews and each review follows the IID Bernoulli distribution, then review valence p' is the percentage of positive reviews (Wu and Ayala Gaytán 2013).

To evaluate the prospect $(V, p; 0, 1 - p)$, expected utility theory uses a mean-variance model of the sample probability p' (e.g., Wu and Ayala Gaytán 2013). Behavioral scholars argue, however, that the mean-variance model does not accurately represent consumer behavior when consumers do not know the true outcome probability, because consumers tend to experience extra uncertainty when dealing with an ambiguous probability. Behavioral theories instead propose nonlinear transformations of the sample probability into a subjective weighting: either a decision weight that measures the desirability of the outcome (Gonzalez and Wu 1999; Tversky and Kahneman 1992) or a subjective probability that measures the subjective likelihood of the outcome (Einhorn and Hogarth 1985; Kahn and Sarin 1988). Following the behavioral theories, we define a consumer's valuation as a weighting process shown below.

$$\begin{aligned} v(V,p';0,\,1-p') &= v(V)w(p',N) + v(0)w(1-p',N) \\ &= v(V)w(p',N), \end{aligned} \qquad (1)$$

where $v(V) > 0$, $v(0) = 0$, and $w(p', N)$ is the consumer's probability weighting function.

The Shape of the Weighting Function in an Online Market

For decisions made under uncertainty, the shape of the probability weighting function $w(p', N)$ represents a consumer's uncertainty preference. Two theoretical frameworks are often adopted in the literature to understand the shapes of the weighting function $w(p', N)$. The classic expected utility theory assumes that a consumer's uncertainty preference (i.e., risk attitude) is *consistent* and independent of outcome probability. The theory further suggests that a consumer's probability weighting function $w(p', N)$ exhibits one of the following three shapes: 1) if a consumer overweights all outcome probabilities, her weighting function is concave (corresponding to risk-seeking behavior). 2) If she underweights all outcome probabilities, her weighting function is convex (corresponding to risk-averse behavior). 3) If she neither overweights nor underweights outcome probabilities, her weighting function is linear (corresponding to risk-neutral behavior).

Behavioral theories like prospect theory (Tversky and Kahneman 1992) state that uncertainty preference can be *inconsistent* based on the probability of an outcome and whether the outcome is a gain or a loss. In other words, a consumer's uncertainty preference can change based on outcome probabilities. Furthermore, Tversky and Kahneman (1992) show that when the outcome is a gain, the majority of people overweight small outcome probabilities but underweight large outcome probabilities, thus creating a reverse *S*-shaped weighting function. Einhorn and Hogarth (Einhorn and Hogarth 1985, 1986; Hogarth and Einhorn 1990) posit that the sample probability may be used as an anchor, and that people tend to adjust the probability upward (i.e., overweighting) if the anchor is a small probability, and adjust the probability downward (i.e., underweighting) if the anchor is a moderate or a large probability. While a reverse *S*-shaped weighting function serves as the general norm (Wakker 2001), both risk-averse behavior for gains of small probabilities and risk-seeking behavior for gains of moderate and large probabilities are observed in the literature (e.g., Tversky and Kahneman 1992). To account for the possibility of these behaviors, especially among consumers with low probability discriminability (Gonzalez and Wu 1999), we also allow for an *S*-shaped weighting function in our framework. Both reverse *S*-shaped and *S*-shaped weighting functions represent consumers' inconsistent uncertainty preferences and have a crossover point at $w(p^*, N) = p^*$ (Einhorn and Hogarth 1985, 1986).

Many factors can influence the magnitude of probability overweighting/underweighting. If a consumer uses a sample probability to estimate the true probability of an outcome, the size of the sample, the source credibility of the sample, and the degree of disagreement among the sources should all influence the magnitude of the subjective probability adjustment (Camerer and Weber 1992; Einhorn and Hogarth 1985, 1986). Larger sample size, higher source credibility, and/or higher consensus among sources lead to a smaller adjustment magnitude and closer proximity of subjective weighting to the sample probability.

We adopt the behavioral framework to conceptualize the weighting function $w(p', N)$ based on the following considerations. First, as discussed above, the behavioral framework is an extension of the expected utility theory because it can accommodate more shapes of the weighting functions. Second, empirical evidence in online markets shows the appropriateness of adopting such framework. For example, due to the newness of online markets, consumers' market experiences are quite heterogeneous. It is found that consumers with greater market experience tend to conform to the predictions of expected utility theory, while those with less market experience are better predicted by prospect theory (List 2004).

Assessing Weighting Function

Researchers use various models to assess probability weighting functions; those models include nonparametric estimation, one-parameter functions and multi-parameter functions (e.g., Gonzalez and Wu 1999; Kahn and Sarin 1988; Schmidt, Starmer, and Sugden 2008). For our research, we adopt and modify Kahn and Sarin's (1988) model and use the following specification to estimate the weighting function:

$$w_i(p', N) = p' - (a_i + b_i p')\frac{p'(1-p')}{N}, \tag{2}$$

where a and b are parameters, and i represents the ith individual.

We choose this model for several reasons. First, the model directly incorporates the variables we are interested in and describes subjective weighting as a function of sample variance $\frac{p'(1-p')}{N}$ (Kahn and Sarin 1988) and sample size N (Einhorn and Hogarth 1985, 1986). Second, the model allows the uncertainty preference $(a_i + b_i p')$ to linearly change with review valence and accommodates all five of the weighting function shapes in our framework. Third, the model allows weighting function shape and crossover point $p^* = -\frac{a_i}{b_i}$ (if b_i is not 0) to be directly estimated by parameters a and b, as shown below.

- When a consumer overweights probabilities, which means $w_i(p', N) > p'$ when $0 < p' < 1$. Parameters a and b can exhibit one of the three relationships: $b_i = 0$ and $a_i < 0$, $b_i > 0$ and $a_i < -b_i$, or $b_i < 0$ and $a_i < 0$. The weighting function has a concave shape.
- When a consumer underweights probabilities, meaning $w_i(p', N) < p'$ when $0 < p' < 1$. Parameters a and b exhibit one of the three relationships: $b_i = 0$ and $a_i > 0$, $b_i > 0$ and $a_i > 0$, or $b_i < 0$ and $a_i > -b_i$. The weighing function has a convex shape.
- When a consumer neither underweights nor overweights probabilities, $w_i(p', N) = p'$. In this case, $a_i = 0$ and $b_i = 0$ and the weighting function has a linear shape.
- When a consumer underweights small probabilities and overweights large probabilities, $w_i(p', N) < p'$ when $0 < p' < p^*$ and $w_i(p', N) > p'$ when $p^* < p' < 1$. So $b_i < 0$ and $0 < a_i < -b_i$. The weighting function has an *S*-shape.
- When a consumer overweights small probabilities and underweights large probabilities, $w_i(p', N) > p'$ when $0 < p' < p^*$ and $w_i(p', N) < p'$ when $p^* < p' < 1$. So $b_i > 0$ and $-b_i < a_i < 0$. The weighting function has a reverse *S*-shape.

The derivations of these results are shown in the Appendix A.

The Impact of Review Volume on WTP

By combining Eqs. (1) and (2), we have

$$\begin{aligned} WTP_i(p', N, V) &= w_i(p', N)v_i(V) \\ &= p'v_i(V) - (a_i + b_i p')\frac{p'(1-p')}{N}v_i(V). \end{aligned} \tag{3}$$

We use Eq. (3) to develop the hypotheses regarding the impact of review volume N on consumer WTP, as shown in the Appendix A. Review volume N is the size of the sample from which review valence p' is obtained, hence a larger review volume results in a smaller magnitude of over/under-weighting. In this study, we focus on the impact of review volume instead of review variance and review valence because 1) review volume and review variance are perfectly negatively correlated if the bidirectional reviews follow a Bernoulli distribution and 2) the impact of review valence on WTP under our conceptual framework is consistent with the literature (e.g., Wu and Ayala Gaytán 2013).

For a seller with a higher review volume (N),

H1. a consumer's WTP is lower if the consumer's $w(p', N)$ is concave.

H2. a consumer's WTP is higher if the consumer's $w(p', N)$ is convex.

H3. a consumer's WTP remains the same if the consumer's $w(p', N)$ is linear.

H4a. a consumer's WTP is higher if the consumer's $w(p', N)$ is *S*-shaped and the review valence (p') is lower than the crossover point $p^* = -\frac{a_i}{b_i}$.

H4b. a consumer's WTP is lower if the consumer's $w(p', N)$ is *S*-shaped and the review valence (p') is higher than the crossover point $p^* = -\frac{a_i}{b_i}$.

H5a. a consumer's WTP is lower if the consumer's $w(p', N)$ is reverse *S*-shaped and the review valence (p') is lower than the crossover point $p^* = -\frac{a_i}{b_i}$.

H5b. a consumer's WTP is higher if the consumer's $w(p', N)$ is reverse *S*-shaped and the review valence (p') is higher than the crossover point $p^* = -\frac{a_i}{b_i}$.

An Experimental Study

To test the internal validity of the framework, we conducted an experimental study with 143 undergraduate business school students from a southern public university.

Study Design and Procedure

We tasked the students with a scenario where they needed to purchase a new LCD TV. As seen in previous studies (e.g., Wu and Ayala Gaytán 2013; Wu et al. 2013), we informed the subjects that they could get the TV from an offline retail store for $800, but instructed them to purchase the TV from an online retail site instead. On this retail site, multiple sellers were offering the same TV with free shipping, and the retail website provided consumer reviews for each individual seller. We showed the subjects a list of sellers with different review profiles and asked them to report the maximum price they were willing to pay each seller for the TV. All seller reviews included in the study were bidirectional in nature, similar to the ones used by eBay. At the beginning of the study, we showed the subjects two examples of review profiles to ensure that all subjects understood the information shown in the study. In keeping with previous research, we employed within-subjects numerical evaluations to assess the individual weighting function (e.g., Tversky and Koehler 1994 used 20 uncertainty stimuli; Einhorn and Hogarth 1985 used 48 uncertainty stimuli). We used a 3 × 11 within-subjects experimental design, with each subject rating 33 seller profiles. Review volume was at one of three pre-defined levels (20, 50, and 200), and review valence was at one of 11 pre-defined levels (0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%). Three profiles (20, 50%; 50, 50%; and 200, 50%) appeared twice in the study to test reliability. We asked the subjects to fully examine all seller profiles before evaluating WTP for each seller, and we allowed them to go back and forth and adjust their WTP freely during the experiment.

We did not use a completely randomized design for two reasons: first, online offers are rarely presented to consumers in a random manner since websites will automatically sort offers by criteria such as review valence and volume. Second, avoiding a completely randomized design reduces the mental burden of comparing many profiles and increased the accuracy of evaluation. To balance the appearance order of review volumes and valences, we developed four versions of the survey for the study. Fig. 1 below is a snapshot of the survey instruments used in the study.

Analysis

Reliability Assessment

We used a Pearson correlation test to assess the test–retest reliability, and removed subjects with a Pearson correlation score of 0.8 and below from the dataset. In the cases where the Pearson correlation test did not apply (e.g., for a subject who reported the same WTP for different sellers), we removed subjects whose preferences for volume reversed between the test and retest sets. 28 of the 143 participants were removed from the data set because they either had missing data or failed the reliability test.

Weighting Function Assessment

We used a linear approximation of Eq. (3) to estimate the weighting function for individual i as follows:

$$WTP = c_i + p'V - (a_i + b_i p')\frac{p'(1-p')}{N}V + \varepsilon_i, \qquad (4)$$

where the constant term c_i was introduced to capture the individual valuation of the perceived benefits from online purchase option besides risk factors (e.g., Weber and Hsee 1998).

Hypothesis Testing

We grouped subjects by the estimated weighting function shape and used the following linear regressions to test our hypotheses at the group level:

$$\mathrm{Log}(WTP_{ijk}) = \alpha_{ijk} + \beta_{pjk}\,\mathrm{Log}(p') + \beta_{Njk}\,\mathrm{Log}(N) + \varepsilon_{ijk}, \qquad (5)$$

where i identifies the ith individual, j represents the jth group, and k denotes below or above the crossover point. We excluded observations with \$0 WTP because in these cases, the subjects did not consider purchasing. Following previous studies (e.g., Ba and Pavlou 2002; Melnik and Alm 2002), we chose a log transformation to capture the diminishing return of reputation on price as seller reputation increased. For the S-shaped and reverse S-shaped groups, we performed separate linear regressions to individually fit the data that fell below and above the crossover point.

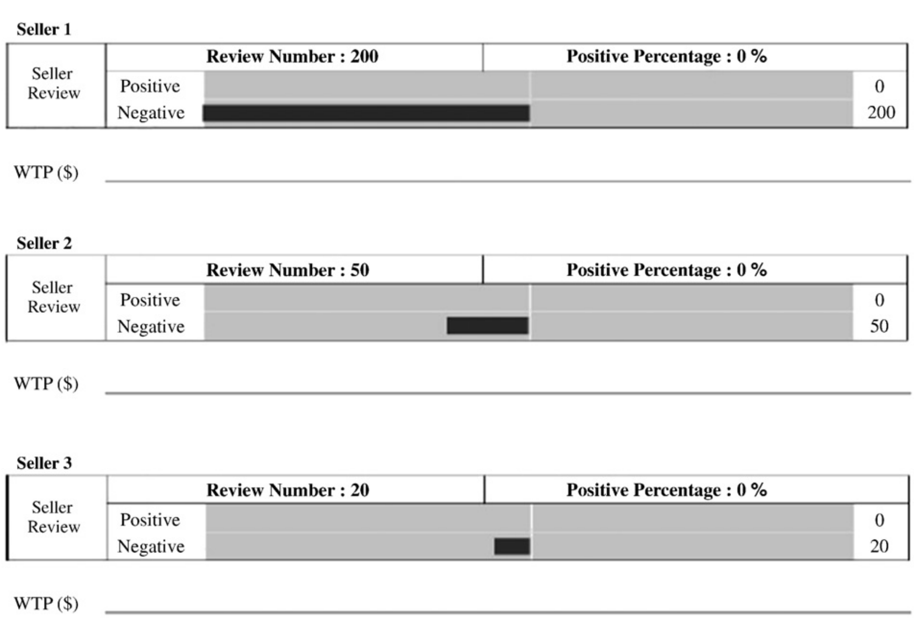

Fig. 1. Experiment study design snapshot.

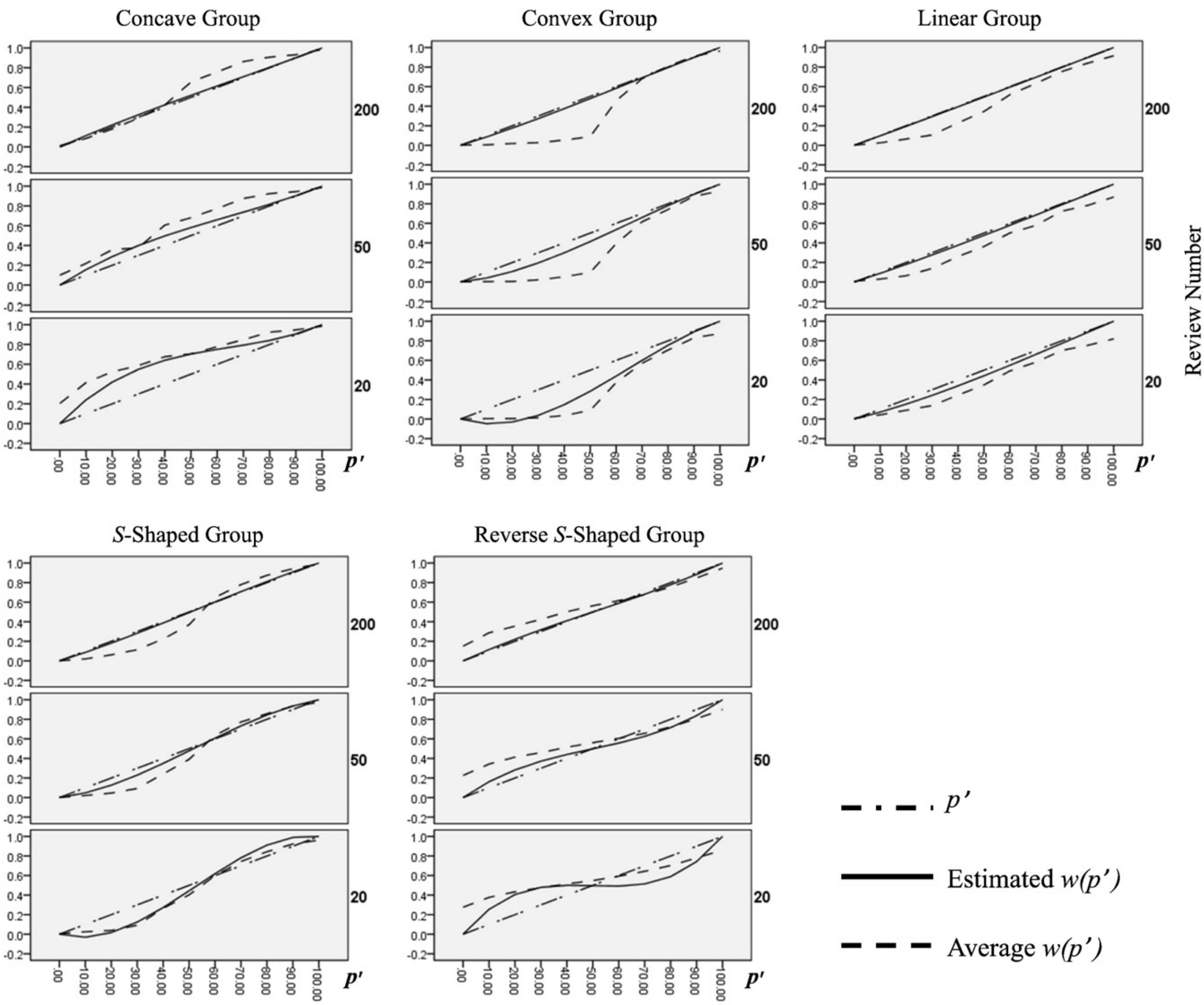

Fig. 2. Plot of weighting functions at group levels.

Results

Weighting Function Assessment

From the data, we were able to identify all five weighting function shapes, which supported our framework. Of all included participants, 39 (34%) consistently overweighted or underweighted all levels of probabilities. Among them, 10 participants had concave-shaped weighting functions, 7 had convex-shaped weighting functions, and 22 had linear-shaped weighting functions. The remaining 76 participants (66%) did not consistently overweight or underweight probabilities. Among them, 25 participants had *S*-shaped weighting functions and 51 participants had reverse *S*-shaped weighting functions. Fig. 2 shows the plots of weighting functions by group.

To make the estimation more balanced between the levels of valence (11 levels) and volume (three levels), we also fitted the model with nine data points where review valance equaled 10%, 50%, and 90%[1]. The classification results changed as expected. For majority of the subjects, the signs of parameter *a* and *b* estimates did not change, but more subjects were classified into the linear group due to a smaller number of observations. Overall, 96% of the subjects were classified as expected in the robustness test.

Hypothesis Testing

Table 1 presents the results of hypothesis testing. Group-level R-squared values ranged from .206 for the linear group to .505 for the *S*-shaped group above the crossover point. As expected, the impact of review volume varied across groups. Supporting H1, H2, and H3, volume negatively influenced WTP ($\beta_N = -1.005$, $p < 0.001$) in the concave group, had a modest positive impact on WTP ($\beta_N = 0.296$, $p < 0.1$) in the

[1] We thank the reviewers for suggesting the robustness checks.

Table 1
The impact of seller review volume on consumers' WTP—experimental study.

Group	Coefficient	Std. error	t-Stat	*p*-Value	R-square	Hypothesis
Concave	−1.005	0.175	−5.748	0.000	0.386	H1 Supported
Convex	0.296	0.155	1.905	0.059	0.498	H2 Supported
Linear	−0.194	0.131	−1.481	0.139	0.260	H3 Supported
S-shaped below cross-over point	−0.027	0.168	−0.158	0.874	0.324	H4a Not supported
S-shaped above cross-over point	−0.184	0.076	−2.440	0.015	0.285	H4b Supported
Reverse *S*-shaped below cross-over point	−0.542	0.081	−6.717	0.000	0.405	H5a Supported
Reverse *S*-shaped above cross-over point	0.107	0.049	2.173	0.030	0.505	H5b Supported

Note: dummy variable is specified for each individual, but estimates are not reported due to large number of parameters for the dummy variables.

convex group, and had no significant effect ($\beta_N = -0.194$, $p > 0.1$) in the linear group. In the *S*-shaped group, volume had no influence on WTP when valence was below the crossover point ($\beta_N = -0.027, p > 0.1$) but had a significant negative impact when valence was above the crossover point ($\beta_N = -0.184$, $p < 0.05$), therefore H4b was supported but H4a was not. Consistent with H5a and H5b, in the reverse *S*-shaped group, volume had a significant negative impact on WTP when valence was below the crossover point ($\beta_N = -0.542$, $p < 0.001$) and a significant positive impact when valence was above the crossover point ($\beta_N = 0.107$, $p < 0.05$).

To check the robustness of the model, we included the observations with $0 WTP and estimated the model again[1]. We found that most of the hypothesis testing results remained the same. The impact of review volume in two sections, the convex group and the reverse *S*-shaped above the cross-over point, became insignificant.

Conclusions from the Experimental Study

Overall, test results support all hypotheses except for H4a and thus, the experimental study provides strong support for our theoretical framework. As a result, we conclude that consumer heterogeneity is more complex than previous studies suggest. The relationship between review volume and WTP varies not only by individual, but also by review valence. Consistent with previous studies (e.g., Chevalier and Mayzlin 2006; Sun 2012; Wu and Ayala Gaytán 2013), review valence had a significant positive impact on WTP in all groups ($\beta_{p'}$ ranged from 0.359 in the reverse *S*-shaped group below the cross-over point to 2.937 in the convex group, all significant at 0.001). Moving forward, we test the external validity of our framework with an empirical study of eBay.com.

An Empirical Study

We conduct the empirical study with two important motivations. First, testing a theoretical framework in online markets is a common approach to establishing its external validity and relevance for managerial implications. Second, a sterile lab setting—as used in the experimental study—may not perfectly reflect the consumer decision process as it naturally occurs in online markets. For example, a majority of online sellers have high review valences. On one hand, studies find that consumers are biased when they process review information, placing more emphasis on review valence and less on review volume (Wolf and Muhanna 2011). On the other hand, the predominance of high review valences diminishes the usefulness of valence in differentiating between good and poor sellers, thus making valence less influential in determining price premium (Bockstedt and Goh 2011).

Data Collection

We select the online auction site eBay.com for use in our empirical study because the prices that consumers bid on products are directly observable on the site and because user reviews are critical for eBay sellers' success. eBay's review system allows a buyer to submit positive, neutral, or negative seller feedback after each transaction, and eBay then aggregates buyers' feedback into two measures: the Feedback Score, which eBay calculates as the number of positive reviews subtracting the number of negative reviews, and the Positive Feedback Percentage (or PFP), which eBay calculates as the number of positive reviews divided by the sum of positive and negative reviews received in the most recent 12 months. eBay calculates both scores for every seller and displays the scores on the auction information page for all potential buyers to view. In addition, buyers can view more detailed review information on each seller by visiting the seller's profile page. This additional information includes the number of positive, neutral, and negative reviews that the seller has received in the past 1, 6, and 12 months.

We collected transaction data for new PlayStation 3 gaming consoles sold on eBay between September and November of 2009. The $299 list price of the PlayStation 3 gaming console held steady during the entire period. From all auctions included in the study, we collected the following data: product description, bidding history, seller profile, and auction information including shipping, payment, and return policy. We collected 678 observations in total, but later removed some observations for various reasons. First, some auctions do not result in sales, yielding invalid transactions. Second, because eBay calculates positive percentage based on reviews in the most recent 12 months, a seller who has negative reviews can still get 100% PFP as long as she receives the negative reviews before the 12-month window. This calculation significantly increases the number of sellers with 100% PFP. To reduce this bias, we documented the 200 most recent reviews for each seller and removed the sellers with 100% PFP despite having negative reviews. Third, some sellers have never before sold an item on eBay but still have 100% positive reviews due to their previous purchase activities. Research shows that reviews for a seller's purchase behavior do not influence sales price (Zhang 2006), so we removed observations corresponding to first-time sellers. Finally, we only kept auctions from US sellers to control for country-of-origin bias (Hu and Wang 2010). After removing 157 observations based on these criteria, we netted a final sample of 529 observations.

Model Variables

eBay adopts a procedure that combines Open English auction and a second-price auction (Kamins, Dreze, and Folkes 2004), which means that the maximum bid of the auction winner is unobservable. On eBay, the bidder who has the highest hidden bid wins the auction and pays a price equal to the second-highest bidder's maximum hidden bid plus a preset increment. Therefore, according to Li, Srinivasan, and Sun (2009), the last observed bidding amount represents the second-highest bidder's maximum bid. Consumers consider shipping cost when participating in auctions, and as such, auctions with higher shipping costs typically result in lower final bidding prices (Bockstedt and Goh 2011). We used the final bidding price plus shipping cost to approximate the second-highest bidder's WTP. We followed the procedure from Kamins, Dreze, and Folkes (2004) to collect the final bidding price of each auction after the auction ended.

eBay calculates a seller's Feedback Score as the difference between the number of positive and negative reviews, so we used the Feedback Score and PFP to approximate a seller's review volume N. We defined review valence p' as the PFP provided by eBay.

Control Variables

We also included other external factors that could affect WTP as control variables. First, eBay lists some featured items at the top of search results, and some items include special features like a warranty. These specialty auctions can influence final bidding price, as consumers may perceive these items as more valuable. Research shows that featured items (Bockstedt and Goh 2011) and items with a full warranty (Zhou, Dresner, and Windle 2009) usually have higher auction prices, so we included a dummy variable named Specialty to differentiate between featured auctions or items with special features (1) and other auctions (0). Next, the ability to return a product reduces the risk associated with a purchase, and as a result, consumers are often willing to pay more for a product if it is returnable. To account for the effect of product return policy, we created a dummy variable Return to indicate whether a product could be returned (0) or not (1). Third, studies find that the number of bidders fully mediates the relationship between the starting bid and the final price (Kamins, Dreze, and Folkes 2004), and that increasing the number of bidders often increases the final bidding price (Suter and Hardesty 2005). To account for this effect, we created a Bidders variable to indicate the number of bidders on each auction.

Auctions ending during peak hours typically have higher closing prices and, in addition, receive more buyer attention during the auction closing period regardless of the length or closing day of the auction (Melnik and Alm 2002). Because of this, we included a dummy variable Hour to indicate the closing period of each transaction. A value 0 indicated an auction ending between 11:00 p.m. and 8:00 a.m. CST, a period that contained an average of 5.3 transactions per hour in our dataset, while a value 1 indicated an auction ending between 8:00 a.m. and 11:00 p.m. CST, a period that showed an average of 32.1 transactions per hour.

Though the list price of the gaming console did not change during the three-month period where we collected the data, the perceived value of the product could change (Park, Bin, and Lee 2011), especially as the holiday season approached. As seen in Wu and Ayala Gaytán (2013), we included two additional dummy variables to account for monthly fluctuations in perceived product value due to external market conditions: dummy variable Month10 identified whether an auction ended in October and dummy variable Month11 identified whether an auction ended in November.

We summarize variables and descriptive statistics of our dataset in Tables 2 and 3, respectively. The average seller-reported shipping cost was $7.88 and the average final bidding price was $294.85, which was lower than the product's list price. Descriptive analysis also indicated that while review volume had a large variance (std. dev. = 1,551.6), review valence was overwhelmingly high with an average of 99.16% and a standard deviation of 0.0163. This distribution of valence is also observed in several previous studies of eBay transactions.

Weighting Function Assessment

We used the same specification (4) with the additional control variables to classify the consumer types:

$$Y_i = c_i + \beta_a X_{ai} + \beta_b X_{bi} + \beta_1 Specialty_i + \beta_2 Return_i + \beta_3 Hour_i + \beta_4 Month_{10i} + \beta_5 Month_{11i} + \beta_6 Bidders_i + \varepsilon_i, \tag{6}$$

Table 2
Summary of empirical data variables.

Variables	Measures
WTP	Final bid plus shipping fee charged by the seller
Review volume N	Total number of reviews
Review valence p'	Percentage of positive reviews
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
Empirical data description.

Variable	Mean	Std. deviation
WTP	302.74	19.16
Review volume N	825.45	1,551.6
Review valence p'	0.9916	0.0163
Bidders	10	4.45
	Number of 0s	Number of 1s
Specialty	482	47
Return	306	223
Hour	48	481
Month 10	343	186
Month 11	383	146

Table 4
Empirical study aggregate results.

Variable	Coefficient	Std. error	t-Stat
Review volume N	0.005	0.001	3.623 *
Review valence p'	0.728	0.146	4.979 *
Specialty	0.016	0.004	4.203 *
Return	−0.005	0.002	−2.194 *
Hour	0.015	0.004	4.083 *
Month 10	−0.002	0.003	−0.867
Month 11	0.014	0.003	5.052 *
Bidders	0.001	0.000	2.781 *
R-square	0.224		

* $p < .05$.

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 is the ith observation.

Since most of the winning bidders had a single observation in our data set, it was impossible to individually estimate consumers' probability weighting functions using model (6). Therefore, we pooled the 529 observations together and estimated model (6) using a finite mixture regression model to obtain segment memberships of each individual (e.g., Wu and Ayala Gaytán 2013). Finite mixture techniques assume that population is formed by multiple subpopulations and thus can be applied to various statistical models in the marketing literature such as regression models, logit models, and structural equation models, depending on the original model structure (e.g., Andrews and Currim 2003; Jedidi, Jagpal, and DeSarbo 1997; Wu and Rangaswamy 2003). Because model (6) is a linear regression model, we thus used a finite mixture regression model of model (6) to classify the 529 observations into segments.

Hypothesis Testing

We replicated the same procedure from the experimental study to test the hypotheses. We estimated model (7) at the segment level based on the segment memberships of the observations; for S-shaped and reverse S-shaped groups, we performed separate analyses to individually fit the data that fell below and above the crossover point.

$$\begin{aligned} \mathrm{Log}WTP_{ijk} = {} & c_{ijk} + \beta_N \,\mathrm{Log}N_{ijk} + \beta_p \,\mathrm{Log}p'_{ijk} + \beta_1 Specialty_{ijk} \\ & + \beta_2 Return_{ijk} + \beta_3 Hour_{ijk} + \beta_4 Month_{10ijk} \\ & + \beta_5 Month_{11ijk} + \beta_6 Bidders_{ijk} + \varepsilon_{ijk}, \end{aligned} \quad (7)$$

where i is the ith observation, j is the jth group, and k is whether the observation is above or below the crossover point.

Results

To benchmark our segment level analysis, we first estimated model (7) with all 529 observations at the aggregate level by assuming the same consumer type in uncertainty preference. The aggregate analysis results are reported in Table 4. At the aggregate level, both review volume and review valence had significant positive influences on consumer WTP. In addition, all control variables except for Month10 significantly impacted WTP. Consistent with previous studies, we found that WTP increased when the number of bidders increased ($\beta_6 = 0.001$, $p < 0.01$), if the auction included a specialty item ($\beta_1 = 0.016$, $p < 0.001$), if the transaction occurred during a peak period ($\beta_3 = 0.015$, $p < 0.001$), and if the seller accepted returns ($\beta_2 = -0.005$, $p < 0.001$). Furthermore, aggregate analysis showed that WTP was higher in November than in September and October ($\beta_5 = 0.014$, $p < 0.001$).

Weighting Function Assessment

We used R package FlexMix to estimate the finite mixture regression model (6) for identifying consumer types. Following Leisch's (2004) approach, we estimated the model with the specified segment number increasing from 1 to 10. For each specified segment number, we ran the analysis 50 times using different random starting values for the parameters, totaling 500 runs. We used Akaide's Information Criterion (AIC) to choose the best model and determine the number of latent segments. In addition, we compared and evaluated models based on another criterion, AIC3—AIC with a per-parameter penalty factor of 3 (Andrews and Currim 2003). Andrews and Currim (2003) tested the performance of seven commonly used segment retention criteria including AIC, AIC3, BIC, CAIC, ICOMP, LOGLV, and NEC for finite mixture regression models. The authors achieved the best overall performance using AIC3, with the highest success rate and a very low parameter bias across various model specifications and data configurations. Compared with BIC, AIC3 was also advantageous when working with small sample sizes. We chose the seven-segment solution that had both the smallest AIC (4,348.188) and AIC3 (4,424.188).

The seven segments identified three out of the five shapes: 20.6% of the consumers belonged to the linear group, 38.2% were S-shaped, and 41.2% were reverse S-shaped. Consistent with both the literature and our experimental study, the reverse S-shaped group was the largest. All observations in the S-shaped group were located above the crossover point, so within the context of our sample, we could consider the S-shaped group to be a convex group. Detailed information for the seven-segment solution is shown in Table 5.

Table 5
Seven-latent-class model parameter estimations.

Latent class	Size	Group	Parameter estimate		Std. error	Z-stat	p-Value	Cross-over point
1	54	Linear	a	6,225.269	5,053.500	1.232	0.218	NA
			b	−6,450.339	5,293.242	−1.219	0.223	
2	55	Linear	a	−610.529	2,268.808	−0.269	0.788	NA
			b	661.048	2,397.805	0.276	0.783	
3	112	Reverse S-shaped	a	−1,296.158	275.667	−4.702	0.000	0.956
			b	1,355.640	293.296	4.622	0.000	
4	59	Reverse S-shaped	a	−2,685.200	93.154	−28.826	0.000	0.941
			b	2,853.300	99.621	28.641	0.000	
5	70	S-shaped	a	636.825	201.779	3.156	0.002	0.891
			b	−714.567	214.832	−3.326	0.001	
6	132	S-shaped	a	1,086.973	305.676	3.556	0.000	0.883
			b	−1,231.051	327.791	−3.756	0.000	
7	47	Reverse S-shaped	a	−1,180.900	95.125	−12.414	0.000	0.938
			b	1,259.600	100.720	12.506	0.000	

Hypothesis Testing

Results showed that the impact of review volume on WTP varied among consumer groups. Consistent with H3, review volume did not affect WTP in the linear group ($\beta_N = 0.004$, $p > 0.1$). In the S-shaped group, review volume showed a negative influence on WTP when review valence was above the crossover point, although the effect was statistically insignificant ($\beta_N = -0.001$, $p > 0.1$). Supporting H5a and H5b, in the reverse S-shaped weighting function group, review volume had a significant negative effect on WTP when the review valence was below the crossover point ($\beta_N = -0.039$, $p < 0.01$) but had a significant positive effect when valence was above that point ($\beta_N = 0.006$, $p < 0.001$). Table 6 presents the results of our hypothesis testing.

Conclusions from the Empirical Study

Although our empirical study identifies only three out of the five shapes originally proposed, it still reveals that consumers differ in their preferences toward review volume. Some consumers simply do not take volume into consideration, some have stable volume preferences, and others change their volume preferences based on review valence. At the aggregate level, our empirical study shows that review volume positively influences consumer WTP because the majority of observations fall in the reverse S-shaped group with review valences above the crossover point. That being said, we expect the relationship between review volume and consumer WTP to change as sample composition changes. Thus, underlying insights about consumer uncertainty preferences are camouflaged and managerial implications are misguided if we draw conclusions about the relationship between user reviews and WTP solely based on aggregate analysis.

Disaggregate analysis showed that review valence had a significant positive impact on WTP except for the linear group ($\beta_p = 0.156$, $p > 0.1$) and the reverse S-shaped group below the crossover point ($\beta_p = -1.631, p < 0.05$). With respect to the linear group, 53 out of 109 observations had a 100% review valence. Even though the impact of review valence was insignificant, the positive coefficient indicated a positive influence on WTP.

Furthermore, aggregate analysis may also obscure the differing impacts of control variables across consumer segments. Specialty had a consistently positive impact on WTP across groups (β_1 ranged from 0.004 in the S-shaped group above the crossover point to 0.040 in the linear group, all significant at 0.1). This result matched our assumption that a specialty item would indicate a greater value to consumers and thus led to a higher price. Additionally, we found that all groups paid a higher price for auctions ending in November (β_5 ranged from 0.005 in the reverse S-shape group above the crossover point to 0.029 in the linear group, all significant at 0.05), which supported our original suggestion that the approaching holiday season would increase consumers' willingness to pay a higher price for the product for various reasons, such as purchasing the product as a gift or paying more for on-time delivery.

Our framework assumed that return policy, which was directly related to purchase uncertainty, should have a varying impact on WTP across groups with different uncertainty preferences. Consistent with the framework, return policy had no influence on WTP for consumers who have no preference regarding uncertainty ($\beta_2 = -0.012$, $p > 0.1$ in linear group) or prefer uncertainty ($\beta_2 = 0.002$, $p > 0.1$ in S-shaped group above the crossover point). For consumers who avoid uncertainty, the ability to return the product to the seller significantly increased

Table 6
The impact of seller review volume on consumers' WTP—empirical study.

Group	Coefficient	Std. error	t-Stat	p-Value	R-square	Hypothesis
Linear	0.004	0.005	0.821	0.414	0.319	H3 Supported
S-shaped above cross-over point	−0.001	0.001	−1.015	0.311	0.447	H4b Not supported
Reverse S-shaped below cross-over point	−0.039	0.008	−5.002	0.007	0.870	H5a Supported
Reverse S-shaped above cross-over point	0.006	0.001	4.364	0.000	0.642	H5b Supported

consumer WTP ($\beta_2 = -0.006$, $p < 0.01$ in reverse S-shaped group below the crossover point).

Discussion

With respect to the impact of online review volume on WTP, there is little consensus in the current literature. We propose that aggregate-level mixed observations may stem from heterogeneity in uncertainty preference among consumers. Both our experimental study and empirical study confirm that such heterogeneity exists and, furthermore, that it influences the ways consumers use online user reviews in their WTP decisions.

Previous studies use cross-consumer heterogeneity to rationalize mixed findings on the impact of online reviews. These researchers acknowledge the interactions among review volume, review variance, and review valence at the aggregate level (Khare, Labrecque, and Asare 2011; Sun 2012) and varying preferences for review volume and review variance at the individual level (Wu and Ayala Gaytán 2013; Wu et al. 2013). When it comes to the interactions between review statistics and the varying preferences within an individual, however, researchers have thus far remained silent. We argue that consumer preferences for review volume and review variance not only differ *across* individuals, but also change with review valence *within* individuals. We show that the interaction between volume and valence can occur at both the individual and aggregate levels. Our framework can be easily used to test other important review characteristics like review variance, review helpfulness, review content, and reviewer credibility (e.g., Chintagunta, Gopinath, and Venkataraman 2010; Nayor, Lamberton, and Norton 2011) by incorporating these factors into uncertainty estimation. Theoretical explanation and the ability to quantify such consumer difference will serve as a foundation for developing segmentation and predicting models that leverage review statistics.

The outcome of our study has important implications for online pricing strategy. Online business has been practicing dynamic pricing based on real-time changes of factors such as competition, purchase timing, and website traffic. Consumer characteristics also come into play in tailoring price toward individuals. Amazon.com has tested dynamic pricing on DVD and MP3 players based on consumers' purchase history and preference in 2000 (Ramasastry 2005). *The Wall Street Journal* conducted research on Staples.com and found that prices were displayed differently to consumers based on their locations and distances to rival stores (Valentino-Devries, Singer-Vine, and Soltani 2012). Our studies on consumer WTP based on review statistics are especially relevant to this issue because "big data allows for a far more scientific approach to selling at different prices, depending on an individual's willingness to pay" (Tanner 2014). While companies need to use dynamic pricing strategies cautiously and must carefully manage potential negative consumer reactions toward price discrimination, we propose that preference regarding review statistics is important information to be incorporated into pricing models at the individual or segment level. We show in our studies that consumers' preferences regarding review statistics are different. When facing sellers with the same review valence, some consumers may pay more to sellers who have larger volumes, while others may pay more to sellers with small volumes. We also show that a consumer's preference regarding review volume may shift: a consumer may be willing to pay more to a seller with a higher volume, but only when valence reaches a certain level. Our framework proposes a method to quantify these characteristics and can be used to more effectively determine the optimal price of a product offered to a consumer.

Our study also has implications for online promotion strategies. As discussed before, review information is shown to increase advertisement click-through rate, and hence the likelihood of purchase. As a result, online advertisements, direct marketing offers, and cross-selling recommendations often include review statistics. Online retailers can customize these promotional activities and display offerings to match a product and/or seller's review profile with a consumer's preference regarding review statistics. New online sellers, who have not established high review statistics, and less successful online sellers, whose review ratings are not the strongest, can better identify potential customers and promote to consumers who have a higher tolerance of sellers with inferior review statistics. Online retailers can also use the same idea to promote second-rate products to the right consumers, or adjust pay-per-click advertising spending based on product/seller review statistics.

The research we present here includes some limitations that help form our recommendations for future studies. First, literature on WTP has suggested various methods to measure consumers' true WTP, including self-explication, incentive-compatible mechanism, Vickery auction, conjoint analysis, etc. Our experiment used self-explicated measurement. Despite the hypothetical bias (Wang, Venkatesh, and Chatterjee 2007), such direct measurement of WTP can still help marketers generate right pricing decisions (Miller et al. 2011). Indeed, we included the 28 subjects who were removed from the experimental study in hypothesis testing and the results did not change dramatically. Regardless, to address the concerns of hypothetical bias, the experimental design can be improved, possibly by employing an incentive-compatible mechanism. Finally, we used point measures of consumer WTP in our analysis. However, WTP may be range-based rather than point-based. We suggest future research to test the framework using different WTP measurements.

Next, the data collection period in our empirical study is not sufficiently long enough to identify sellers with low review valences because the market may eliminate such sellers before we are able to collect the data. As a result, we collected fewer observations below the crossover points in the S-shaped and the reverse S-shaped consumer groups. Moreover, a similar issue exists with regard to product user reviews because online retailers often purposefully remove products with bad reviews (Banjo 2012). The results can be further validated by future studies that employ larger and more representative sample datasets.

Third, our studies establish correlation, but not causality, when it comes to the relationship between weighting functions and the impact of review volume on WTP. Going forward, researchers should establish causality by developing independent measurements of weighting functions. Finally, we develop our current framework using the binary review format, and thus it only takes review volume and review valence into consideration. Future research can extend our framework to accommodate continuous review formats (e.g., Amazon.com's five-star review system) that draw upon review variance, in addition to volume and valence, as they influence consumer purchase decisions. Testing whether a consumer's weighting function of review statistics differs for search and experience products would also be interesting, as Jimenez and Mendoza (2013) indicate that consumers may process online reviews differently for search and experience products.

Appendix A

A.1. The Shape of Weighting Function and Parameters a and b

The weighting function is determined by Eq. (2) in the paper

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

1. A consumer overweights all p' when $0 < p' < 1$, the weighting function has a concave shape

   $$p' - (a_i + b_i p')\frac{p'(1-p')}{N} > p'.$$

   Since $\frac{p'(1-p')}{N} > 0$, which means $a_i < -b_i p'$ for all p' when $0 < p' < 1$
   1) When $b_i = 0$ and $a_i < 0$,
   2) When $b_i > 0$ and $a_i < -b_i$,
   3) When $b_i < 0$ and $a_i < 0$.
2. A consumer underweights all p' When $0 < p' < 1$, the weighting function has a convex shape

   $$p' - (a_i + b_i p')\frac{p'(1-p')}{N} < p'.$$

   Since $\frac{p'(1-p')}{N} > 0$, which means $a_i > -b_i p'$ for all p' when $0 < p' < 1$
   1) When $b_i = 0$ and $a_i > 0$,
   2) When $b_i > 0$ and $a_i > 0$,
   3) When $b_i < 0$ and $a_i > -b_i$.
3. A consumer neither underweights nor overweights all p' When $0 < p' < 1$, the weighting function has a linear shape

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

   Since $\frac{p'(1-p')}{N} > 0$, which means $a_i = -b_i p'$ for all p' when $0 < p' < 1$, so $a_i = 0$ and $b_i = 0$.
4. A consumer underweights all p' when $0 < p' < p^*$ and overweights all p' when $p^* < p' < 1$, where $0 < p^* = -\frac{a_i}{b_i} < 1$; the weighting function has a S-shape.
   Because $0 < -\frac{a_i}{b_i} < 1$, so
   1) When $b_i > 0$, $-b_i < a_i < 0$
   2) When $b_i < 0$, $0 < a_i < -b_i$.

Consider $0 < p_1' < -\frac{a_i}{b_i} < p_2' < 1$.
Because $w(p_1') < p_1'$ for all p_1' and $w(p_2') > p_2'$ for all p_2', so $a_i + b_i p_1' > 0$ and $a_i + b_i p_2' < 0$, which means $-b_i p_1' < a_i < -b_i p_2'$.
Because $p_1' < p_2'$, so $b_i < 0$.
As a result, $b_i < 0$ and $0 < a_i < -b_i$.

5. A consumer overweights all p' when $0 < p' < p^*$ and underweights all p' when $p^* < p' < 1$, where $w(p^*, N) = p^* = -\frac{a_i}{b_i}$; the weighting function has a S-shape.
   Because $0 < -\frac{a_i}{b_i} < 1$, so
   1) When $b_i > 0$, $-b_i < a_i < 0$
   2) When $b_i < 0$, $0 < a_i < -b_i$.

Consider $0 < p_1' < -\frac{a_i}{b_i} < p_2' < 1$.
Because $w(p_1') > p_1'$ for all p_1' and $w(p_2') < p_2'$ for all p_2', so $a_i + b_i p_1' < 0$ and $a_i + b_i p_2' > 0$, which means $-b_i p_2' < a_i < -b_i p_1'$.
Because $p_1' < p_2'$, so $b_i > 0$.
As a result, $b_i > 0$ and $-b_i < a_i < 0$.

A.2. The Impact of Review Volume N on WTP

By Eq. (3), we have

$$\frac{\partial WTP_i\,(p', N,\ V)}{\partial N} = \frac{v_i(V)}{N^2}(a_i + b_i p')\left(p' - p'^2\right) = \frac{v(V)}{N^2}(a_i + b_i p')p'(1-p').$$

H1. For consumers with concave weighting functions, parameters a and b can exhibit one of the three relationships: $b_i = 0$ and $a_i < 0$, $b_i > 0$ and $a_i < -b_i$, $b_i < 0$ and $a_i < 0$.

Because $v(V) \geq 0$, when $0 < p' < 1$, $(1 - p') > 0$.

1) When $b_i = 0$ and $a_i < 0$,

$$\frac{\partial WTP_i\,(p', N, V)}{\partial N} = \frac{v_i(V)}{N^2}a_i p'(1-p') < 0.$$

2) When $b_i > 0$ and $a_i < -b_i$,

$$\frac{\partial WTP_i\,(p', N, V)}{\partial N} = \frac{v_i(V)}{N^2}(a_i + b_i p')p'^{(1-p')} < \frac{v_i(V)}{N^2}(-b_i + b_i p')p'(1-p') < -\left[\frac{v_i(V)}{N^2}b_i p'(1-p')^2\right] < 0.$$

3) When $b_i < 0$ and $a_i < 0$,

$$\frac{\partial WTP_i\,(p', N, V)}{\partial N} = \frac{v_i(V)}{N^2}(a_i + b_i p')p'(1-P') < 0.$$

Therefore, Review volume N has a negative impact on WTP.

H2. For consumers with convex weighting functions, parameters a and b exhibit one of the three relationships: $b_i = 0$ and $a_i > 0$, $b_i > 0$ and $a_i > 0$, or $b_i < 0$ and $a_i > -b_i$.

1) When $b_i = 0$ and $a_i > 0$,

$$\frac{\partial WTP_i\,(p', N, V)}{\partial N} = \frac{v(V)}{N^2} a_i p'(1-p') > 0.$$

2) When $b_i > 0$ and $a_i > 0$,

$$\frac{\partial WTP_i(p', N, V)}{\partial N} = \frac{v_i(V)}{N^2}(a_i + b_i p')p'^{(1-p')} > 0.$$

3) $b_i < 0$ and $a_i > -b_i$,

$$\begin{aligned}\frac{\partial WTP_i(p', N, V)}{\partial N} &= \frac{v_i(V)}{N^2}(a_i + b_i p')p'(1-p') \\ &> \frac{v_i(V)}{N^2}(-b_i + b_i p')p'(1-p') \\ &> -\left[\frac{v_i(V)}{N^2} b_i p'(1-p')^2\right] > 0.\end{aligned}$$

Therefore, Review volume N has a positive impact on WTP.

H3. For consumers with linear weighting functions, $a_i = 0$ and $b_i = 0$.

$$\frac{\partial WTP_i(p', N, V)}{\partial N} = \frac{v_i(V)}{N^2}(a_i + b_i p')p'(1-p') = 0$$

Review volume N has no impact on WTP.

H4. For consumers with S-shaped weighting functions, $b_i < 0$ and $0 < a_i < -b_i$.

H4a. When $0 < p' < -\frac{a_i}{b_i}$, $-a_i < b_i p' < 0$.

$$\begin{aligned}\frac{v_i(V)}{N^2}(a_i - a_i)p'(1-p') &< \frac{\partial WTP_i\,(p', N, V)}{\partial N} \\ &< \frac{v_i(V)}{N^2} a_i p'(1-p') \frac{\partial WTP_i\,(p', N, V)}{\partial N} > 0\end{aligned}$$

Review volume N has a positive impact on WTP.

H4b. When $-\frac{a_i}{b_i} < p' < 1$, $b_i < b_i p' < -a_i$.

$$\begin{aligned}2\frac{v_i(V)}{N^2} b_i p'(1-p') &< \frac{\partial WTP_i\,(p', N, V)}{\partial N} \\ &< \frac{v_i(V)}{N^2}(a_i - a_i)p'(1-p') \frac{\partial WTP_i(p', N, V)}{\partial N} < 0\end{aligned}$$

Review volume N has a negative impact on WTP.

H5. For consumers with reverse S-shaped weighting functions, $b_i > 0$ and $-b_i < a_i < 0$.

H5a. When $0 < p' < -\frac{a_i}{b_i}$, $0 < b_i p' < -a_i$.

$$\begin{aligned}\frac{v_i(V)}{N^2} a_i p'(1-p') &< \frac{\partial WTP_i\,(p', N, V)}{\partial N} \\ &< \frac{v_i(V)}{N^2}(a_i - a_i)p'(1-p') \frac{\partial WTP_i(p', N, V)}{\partial N} < 0\end{aligned}$$

Review volume N has a negative impact on WTP.

H5b. When $-\frac{a_i}{b_i} < p' < 1$, $-a_i < b_i p' < b_i$.

$$\begin{aligned}\frac{v_i(V)}{N^2}(a_i - a_i)p'(1-p') &< \frac{\partial WTP_i(p', N, V)}{\partial N} \\ &< 2\frac{v_i(V)}{N^2} b_i p'(1-p') \frac{\partial WTP_i(p', N, V)}{\partial N} > 0\end{aligned}$$

Review volume N has a positive impact on WTP.

References

Anderson, Myles (2014), "Local Consumer Review Survey 2014," http://www.brightlocal.com/2014/07/01/local-consumer-review-survey-2014/ (retrieved on January 10, 2015).

Andrews, Rick L. and Imran S. Currim (2003), "Retention of Latent Segments in Regression-based Marketing Models," *International Journal of Research in Marketing*, 20, 4, 315–21.

Ba, Sulin and Paul A. Pavlou (2002), "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly*, 26, 3, 243–68.

Banjo, Shelly (2012), "Firms Take Online Reviews to Heart," *The Wall Street Journal* (http://online.wsj.com/news/articles/SB10001424052702303292204577517394043189230, retrieved on October 27, 2013).

Basuroy, Suman, Subimal Chatterjee, and S. Abraham Ravid (2003), "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," *Journal of Marketing*, 67, 4, 103–17.

Bockstedt, Jesse and Kim Huat Goh (2011), "Seller Strategies for Differentiation in Highly Competitive Online Auction Markets," *Journal of Management Information Systems*, 28, 3, 235–67.

Camerer, Colin and Martin Weber (1992), "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 5, 4, 325–70.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43, 3, 345–54.

Chintagunta, Pradeep K., Shyam Gopinath, and Sriram Venkataraman (2010), "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets," *Marketing Science*, 29, 5, 944–57.

Clemons, Eric K., Guodong Gao, and Lorin M. Hitt (2006), "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems*, 23, 2, 149–71.

Einhorn, Hillel J. and Robin M. Hogarth (1985), "Ambiguity and Uncertainty in Probabilistic Inference," *Psychological Review*, 92, 4, 433–61.

——— and ——— (1986), "Decision Making Under Ambiguity," *The Journal of Business*, 59, 4, S225–50.

Friedman, Dan (2011), "5 Simple Ways to Improve your Ad Words Performance," http://adwords.blogspot.com/2011/04/5-simple-ways-to-improve-your-adwords.html (retrieved on January 10, 2015).

Gonzalez, Richard and George Wu (1999), "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 1, 129–66.

Grewal, Dhruv, Ramkumar Janakiraman, Kalyanam Kirthi, P.K. Kannan, Brian Ratchford, Reo Song, and Stephen Tolerico (2010), "Strategic Online and Offline Retail Pricing: A Review and Research Agenda," *Journal of Interactive Marketing*, 24, 2, 138–54.

Hogarth, Robin M. and Hillel J. Einhorn (1990), "Venture Theory: A Model of Decision Weights," *Management Science*, 36, 7, 780–803.

Holloway, Darnell (2011), "How Are Yelp's Search Results Ranked?" https://biz.yelp.com/blog/how-are-yelps-search-results-ranked (retrieved on January 10, 2015).

Hu, Ye and Xin Wang (2010), "Country-of-Origin Premiums for Retailers in International Trades: Evidence from eBay's International Markets," *Journal of Retailing*, 86, 2, 200–7.

Jang, Sunghua, Ashutosh Prasad, and Brian T. Ratchford (2012), "How Consumers Use Product Reviews in the Purchase Decision Process," *Marketing Letters*, 23, 3, 825–38.

Jedidi, Kamel, Harsharanjeet S. Jagpal, and Wayne S. DeSarbo (1997), "Finite-Mixture Structural Equation Models for Response-based Segmentation and Unobserved Heterogeneity," *Marketing Science*, 16, 1, 39–59.

Jimenez, Fernando R. and Norma A. Mendoza (2013), "Too Popular to Ignore: The Influence of Online Reviews on Purchase Intentions of Search and Experience Products," *Journal of Interactive Marketing*, 27, 3, 226–35.

Kahn, Barbara E. and Rakesh K. Sarin (1988), "Modeling Ambiguity in Decisions Under Uncertainty," *Journal of Consumer Research*, 15, 2, 265–72.

Kamins, Michael A., Xavier Dreze, and Valerie S. Folkes (2004), "Effects of Seller-Supplied Prices on Buyers' Product Evaluations: Reference Prices in an Internet Auction Context," *Journal of Consumer Research*, 30, 4, 622–8.

Khare, Adwait, Lauren I. Labrecque, and Anthony K. Asare (2011), "The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings," *Journal of Retailing*, 87, 1, 111–26.

King, Robert Allen, Pardeep Racherla, and Victoria D. Bush (2014), "What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature," *Journal of Interactive Marketing*, 28, 3, 167–83.

Leisch, Friedrich (2004), "FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R," *Journal of Statistical Software*, 11, 8.

Li, Shino, Kannan Srinivasan, and Baohong Sun (2009), "Internet Auction Features as Quality Signals," *Journal of Marketing*, 73, 1, 75–292.

List, John A. (2004), "Neoclassical Theory Versus Prospect Theory: Evidence from The Marketplace," *Econometrica*, 72, 2, 615–25.

Melnik, Mikhail I. and James Alm (2002), "Does a Seller's eCommerce Reputation Matter? Evidence from eBay Auctions," *The Journal of Industrial Economics*, 50, 3, 337–49.

Miller, Klaus M., Reto Hofstetter, Harley Krohmer, and Z. John Zhang (2011), "How Should Consumers' Willingness to Pay be Measured? An Empirical Comparison of State-of-the-Art Approaches," *Journal of Marketing Research*, 48, 1, 172–84.

Moe, Wendy W. and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48, 3, 444–56.

Nayor, Rebecca, Cait Lamberton, and David Norton (2011), "Seeing Ourselves in Others: Reviewers Ambiguity, Egocentric Anchoring, and Persuasion," *Journal of Marketing Research*, 48, 3, 617–31.

Nielsen (2010), "Global Trends in Online Shopping, a Nielsen Global Consumer Report," www.hk.hielsen.com/documents/Q12010OnlineShoppingTrendsReport.pd (retrieved on October 27, 2013).

——— (2012), "Global Trust in Advertising and Brand Messages," www.fi.nielsen.com/site/documents/NielsenTrustinAdvertisingGlobalReportApril2012.pdf (retrieved on October 27, 2013).

Park, JaeHong, Gu Bin, 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*, 11, 3, 253–61.

Ramasastry, Anita (2005), "Websites Change Prices Based on Customers' Habits," http://www.cnn.com/2005/LAW/06/24/ramasastry.website.prices/ (retrieved on January 18, 2015).

Schmidt, Ulrich, Chris Starmer, and Robert Sugden (2008), "Third-Generation Prospect Theory," *Journal of Risk and Uncertainty*, 36, 3, 203–23.

Simonson, Itamar and Emanuel Rosen (2014), "What Marketers Misunderstand About Online Reviews," *Harvard Business Review*, 92, 1/2, 23–5.

Sun, Monic (2012), "How Does the Variance of Product Ratings Matter?" *Management Science*, 58, 4, 696–707.

Suter, Tracy A. and David M. Hardesty (2005), "Maximizing Earnings and Price Fairness Perceptions in Online Consumer-to-Consumer Auctions," *Journal of Retailing*, 81, 4, 307–17.

Tanner, Adam (2014), "Different Customers, Different Prices, Thanks to Big Data," *Forbes*, http://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/ (retrieved on January 18, 2015).

Tversky, Amos and Daniel Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 4, 297–323.

——— and Derek J. Koehler (1994), "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101, 4, 547–67.

Valentino-Devries, Jennifer, Jeremy Singer-Vine, and Ashkan Soltani (2012), "Websites Vary Prices, Deals Based on Users' Information," *The Wall Street Journal* (http://www.wsj.com/news/articles/SB10001424127887323777204578189391813881534, retrieved on January 18, 2015).

Wakker, Peter P. (2001), "Testing and Characterizing Properties of Nonadditive Measures Through Violations of the Sure-Thing Principle," *Econometrica*, 69, 4, 1039–59.

Wang, Tuo, Ramaswamy Venkatesh, and Rabikar Chatterjee (2007), "Reservation Price as a Range: An Incentive-Compatible Measurement Approach," *Journal of Marketing Research*, 44, 2, 200–13.

Weber, Elke U. and Christopher Hsee (1998), "Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk," *Management Science*, 44, 9, 1205–17.

Wolf, James R. and Waleed A. Muhanna (2011), "Feedback Mechanisms, Judgment Bias, and Trust Formation in Online Auctions," *Decision Sciences*, 42, 1, 43–68.

Wu, Jianan and Arvind Rangaswamy (2003), "A Fuzzy Set Model of Search and Consideration with an Application to an Online Market," *Marketing Science*, 22, 3, 411–34.

——— and Edgardo Arturo Ayala Gaytán (2013), "The Role of Online Seller Reviews and Product Price on Buyers' Willingness-to-Pay: A Risk Perspective," *European Journal of Information Systems*, 22, 4, 416–33.

———, Yinglu Wu, Jie Sun, and Zhilin Yang (2013), "User Reviews and Uncertainty Assessment: A Two Stage Model of Consumers' Willingness-to-Pay in Online Markets," *Decision Support Systems*, 55, 1, 175–85.

Zhang, Jie (2006), "The Roles of Players and Reputation: Evidence from eBay Online Auctions," *Decision Support Systems*, 42, 3, 1800–18.

Zhou, Ming, Martin Dresner, and Robert Windle (2009), "Revisiting Feedback Systems: Trust Building in Digital Markets," *Information Management*, 46, 5, 279–84.

Zhu, Feng and Xiaoquan (Michael) Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74, 2, 133–48.

Yinglu Wu is an Assistant Professor of Marketing at Boler School of Business, John Carroll University. She earned her Ph.D. of Business Administration from E. J. Ourso College of Business, Louisiana State University in 2012. Her research interests include e-commerce and consumer online purchase behavior. Her research has been published in *Decision Support Systems*.

Jianan Wu is an Associate Professor of Marketing at LSU's E. J. Ourso College of Business. He earned his Ph.D. of Business Administration from the Smeal College of Business, the Pennsylvania State University in 1998. His research interests are in the areas of customer behavior and firm strategy in electronic markets. His research has been published in leading academic journals such as *Journal of Marketing Research, Marketing Science, and Information Systems Research* etc. and has earned international research awards (e.g., the IJRM-EMAC Steenkamp Award 2011). He has served on the editorial review board of *Journal of Interactive Marketing* since 2007.