Marketing Science Institute Working Paper Series 2018
Report No. 18-106

The Fateful First Consumer Review

Sungsik Park, Woochoel Shin, and Jinhong Xie

Report Summary

While extensive research has demonstrated myriad benefits of user-generated product information, in this study Sungsik Park, Woochoel Shin, and Jinhong Xie identify a crucial weakness of such information—"information availability" bias. They provide theoretical and empirical evidence that this weakness leads to a surprising "first-review effect," under which a product's first consumer review has the power to influence subsequent reviews (and the fate of the product) long into the future.

They demonstrate the significance of this first-review effect using data from multiple product categories and multiple platforms. For example, more than 30 percent of vacuum cleaner models simultaneously offered by both Amazon and Best Buy receive first reviews of opposite valence on the two platforms. Both the average review rating and the number of reviews are significantly lower on the platform with a negative first review than on their counterparts with a positive first review (on average, 0.60 fewer stars and 25 percent fewer review postings, respectively).

More strikingly, a negative first review harms the average review rating even after 36 months, and the damage of a negative first review on the number of reviews increases, rather than decreases, over time.

The data from Amazon also show that the first review rating is *not* correlated with product quality, which implies that a high-quality product has a considerable chance of receiving a negative first review. As a result, a single review may destroy a good product's chance of market success, which injures all parties involved: the seller, the platform, and consumers.

The surprisingly persistent and increasing first-review effect is fundamentally driven by a basic property of consumer reviews: without sales, there can be no reviews. Consequently, when a product receives an unfavorable first review, it not only suffers low initial sales but also loses the opportunity to generate a viable number of reviews in the future and thus, the opportunity to correct any potential negative bias of an initial review via subsequent reviews. This information-availability bias forms a mechanism to transfer a disadvantage from a product's first review to a long-lasting and even increasing disadvantage in future word-of-mouth information.

In an era where management's attention is increasingly turning to big data, this study demonstrates the influence of a single data point (the first review) on product success. Given the significant power of this first review effect, firms will need to develop strategies to manage it. The authors offer the following suggestions:

For manufacturers:

- Vigilance: Closely monitor online platforms to detect when a product's first review is posted on each.
- Quick response: Take action to facilitate WOM as soon as an unfavorable first review appears.
- Encourage early reviews: (e.g., participate in Amazon's Early Reviewer Program or Vine Program).

- Integrate communication strategy: Recognize the information-availability bias of user-generated information and integrate firm- and user-generated information into an effective communication strategy.

For online sellers/platforms:

- Facilitate early reviews (e.g., offer a platform-initiated review incentive program, such as Amazon's Early Reviewer Program or Vine Program).
- Link to consumer review metasites (such as ConsumerReview.org) that combine consumer review information from multiple sources).

Sungsik Park is a doctoral student, Woochoel Shin is Associate Professor, and Jinhong Xie is Professor and J.C. Penney Eminent Scholar Chair, all at the Warrington College of Business, University of Florida.

1 Introduction

In today's markets, consumers increasingly depend on online word of mouth (WOM) as a reliable source of information in their purchase decisions (eMarketer 2010). According to recent consumer surveys, 90 percent of consumers report that their buying decisions are influenced by online consumer reviews (Dimensional Research 2013), and 88 percent of consumers trust online reviews as much as personal recommendations (BrightLocal, 2014). Motivated by the practical significance, academic interest in online consumer reviews has surged and the literature provides extensive evidence to support the impact of online reviews on consumer purchasing behavior (e.g., Ameri et al. 2016, Chen and Xie 2008, Chintagunta et al. 2010, Godes and Mayzlin 2004, Kuksov and Xie 2010, Liu 2006, Sun 2012). Recently, three meta-analyses concluded that the two key metrics of online WOM, valence and volume, are directly linked to product sales (Babić et al. 2016, Floyd et al. 2014, You et al. 2015).

Given the strong linkage between online WOM and sales, the importance of obtaining favorable consumer reviews to product success cannot be overemphasized. The literature has suggested that consumer reviews are affected by product-specific characteristics, such as functionality, brand image, and price (e.g., Chen et al. 2011, De Langhe et al. 2016, Godes and Silva 2012, Li and Hitt 2008 and, Luca and Zervas 2016). Thus, if the same product is available at a similar price on multiple online retail platforms, its online reviews should be similar across platforms. However, many market observations contradict this expectation, as shown in Table 1 (Tables follow Reference throughout).

All three products listed in Table 1 are available at both amazon.com and walmart.com.[1] Although both platforms use the same five-star evaluation scale, online reviews of these products are markedly different across these two platforms. First, consider Product A, offered at the same price on both platforms. Its average review rating is 4 stars at Amazon, but only 2.2 stars at Walmart, with 303 reviews at Amazon but only four reviews at Walmart. Second, even in instances where different platforms charge different prices, consumer reviews vary across the platforms, although not necessarily because of the cross-platform price differential. For example, both Products B and C are sold at a lower price at Amazon than at Walmart. However, Amazon's lower price is a disadvantage in online WOM for Product B, but an advantage in online WOM for Product C. Such cross-platform inconsistency in consumer reviews is not limited to these three products or

[1] For brevity, we will refer to these online platforms as "Amazon" and "Walmart" respectively from this point on.

between the two websites presented in Table 1. Rather, it exists in many product categories across various online platforms.

These observations pose several questions. Why does a product with favorable reviews on one online platform receive an unfavorable review on another? Is this a more general phenomenon than we have so far realized? What fundamental market forces drive such substantial cross-platform inconsistency? Is it possible to predict the direction of such inconsistencies in online WOM? What are the important implications to firms' online WOM management policies, especially given the significance of online WOM to a product's success? Answers to these questions are of significance both from theoretical and practical perspectives.

To answer these questions, this paper investigates the evolution of online consumer reviews, both theoretically and empirically. We begin by developing a theoretical model that characterizes the process of online WOM evolution. By explicitly modeling the influence of online reviews on consumer purchase decisions, as well as the process of updating online reviews, we discover that the two key metrics of online WOM, valence and volume, are not independent; rather, they evolve interdependently following a positive or negative feedback pattern. Specifically, our analysis reveals that, at any given time, the current WOM valence positively impacts future WOM volume, and the current WOM volume positively or negatively impacts future WOM valence, depending on whether the current WOM valence has an upward or a downward bias.

We show that such interdependence between WOM valence and volume is fundamentally driven by the distinctive characteristics of consumer reviews as a unique information source. Unlike other types of product information (such as advertising and third-party product reviews), online consumer reviews are, by definition, posted by users based on their personal consumption experience, which implies that the availability of consumer reviews for a given product is *conditional* on its adoption. First, this conditional availability of consumer reviews creates the dependence of WOM volume on valence. In other words, at any given time, an increase in a product's average review rating leads to more sales, which, in turn, boosts its future review-posting volume. Second, though less intuitive, this conditional availability of consumer reviews also creates the dependence of WOM valence on volume through the updating process of WOM valence. Specifically, at any given time, a product's average review rating is updated as the weighted average of two components: 1) the previous average rating, which is based on previously posted reviews, and 2) the incremental average rating, which is based on newly posted reviews.

The larger the volume of the newly posted reviews, the stronger is their impact on the updated valence.

Understanding the interdependence between WOM valence and volume is vital because, as will be shown by our theoretical analysis, this interdependence endows a product's first consumer review with striking power. In today's information-rich environment, where online shoppers have a staggering amount of user-generated information detailing other shoppers' consumption experience, it is hard to believe that a single consumer review could exert a significant influence over the fate of a given product. We find, however, that a product's first consumer review has a *persistent* impact on the product's entire WOM history, measured by both volume and valence. More surprisingly, the impact of the first review on future WOM volume does not diminish, but rather, it intensifies over time. Our theoretical analysis leads to four specific predictions regarding the first-review effect, as summarized in Table 2 (see A1~A4). The first two predictions focus on the first review's overall effect on WOM valence and volume, and the last two predictions focus on the first review's dynamic effect on WOM valence and volume.

These predictions suggest a persistent and even increasing first-review effect, which is counterintuitive and seemingly unlikely since a single review, however persuasive it may be, is likely to be buried down in the list as subsequent reviews are posted. Thus, intuitively, other reviews should eventually cancel out the impact of the initial review. To gain some external validity, we test these theoretical predictions based on field data. Specifically, we conduct an empirical study using a sample of 177 vacuum cleaner models sold simultaneously on two different retail platforms, Amazon and Best Buy. We empirically examine the relationships between a product's cross-platform difference in its first review and its cross-platform difference in future WOM valence and volume by developing an econometric model that controls other cross-platform differences (such as differences in price) and platform-specific effects. Consistent with A1 and A2, our results reveal that, when the same product is offered on the two platforms, both the average review rating and the number of reviews are significantly lower on the platform with a negative first review than on that with a positive first review (on average, 0.60 fewer stars and 25.35 percent less review postings, respectively).[2] Our results also support the dynamic patterns of the first-review effect predicted in A3 and A4. The first-review effect on WOM valence becomes weaker

[2] In the examples given in Table 1, their cross-platform (dis)advantage in WOM valence and volume are also associated with their cross-platform (dis)advantage in the first-review rating.

as time passes, but it does not disappear even after an extended period of time. Specifically, our estimates suggest that a negative first review decreases the average rating by 0.19 stars, even one year after the first review posting. Our results also support the counterintuitive snowball effect of the first review on WOM volume (i.e., the impact of the first review on the number of reviews increases with time). Specifically, our estimates reveal that, on average, a product with a negative first review receives 10.72 fewer reviews at the end of the 6^{th} month and 43.93 fewer reviews at the end of the 12^{th} month than a product with a positive first review.

We supplement our empirical analysis by various robustness checks, which generate the following results. First, the first-review effects hold for different lengths of observation window. Specifically, using the data set of one-, two-, and three-year observation windows, we can confirm that the first-review effect persists even after a significant period of time. More surprisingly, the snowball effect is even supported under an observation window as long as three years. Second, the first-review effect survives some potential alternative explanations. In particular, our analysis shows that the first-review effect is unlikely to be driven by distinct tastes of consumers in different platforms towards certain unobserved product characteristics. Moreover, even after accounting for an increase in volume resulting from a front-page placement, we still observe the first-review effect on WOM volume. Third, the first-review effect is not specific to a certain product category or platform, but is a general finding across various contexts. Our analysis reveals that the data from another category (toasters) support all of our predictions, and so do those from another pair of platforms (Amazon US vs. Amazon Canada). Finally, the first-review effect is robust to alternative assumptions on price variations and platform heterogeneity. More specifically, we obtain the same results even after accounting for temporal price variation. Furthermore, even when the magnitude of the first-review effect is assumed to be different across platforms, we still observe the persistence of the first-review effect.

Taken together, our empirical studies provide evidence that supports our theoretical predictions and demonstrates the formidable and protracted influence of a product's first review on its future online WOM valence and volume.

This paper contributes to online WOM research and practice. In recent years, scholars have made significant efforts to study the consequences of online WOM. In particular, many papers have investigated the influence of online WOM on sales (e.g., Chen et al. 2011, Chintagunta et al. 2010, Libai et al. 2013, Liu 2006, Ludwig et al. 2013 and Trusov et al. 2009), as well as the boundary conditions for such influences (e.g., Forman et al. 2008, Ho-Dac et al. 2013 and Zhu and Zhang 2010). Based on the results of a large number of empirical studies, recent meta-analyses concluded that two key metrics of WOM, volume and valence, significantly affect sales (Babić et al. 2016, Floyd et al. 2014, You et al. 2015). In this paper, we uncover the interdependence between these two key metrics of online WOM and derive a crucial implication of such interdependence: the powerful first-review effect. Our research findings make the following contributions.

First, while "big data" is one of the hottest buzz-phrases in today's business world, our research brings attention to the significance of "small" data. The first review, even though it is a single data point, has the potential to sway the entire evolution path of online consumer reviews and, thus, the fate of the product. As will become clear in our theoretical model, the power of the first review is created by the interdependence between WOM valence and volume. Fundamentally, this interdependence forms a mechanism that transfers a (dis)advantage in a product's first review rating to (1) a *long-lasting* (dis)advantage in future WOM valence, and (2) an *increasing* (dis)advantage in future WOM volume. Therefore, the first review, although small in size, holds significant predictive power. Furthermore, our data surprisingly reveal that the valence of the first review is not correlated with product quality, implying that a high-quality product has the same odds of receiving a negative first review as does a low-quality product. These findings call for proactive strategies in managing the first review. In the early stages of a product's life cycle, a special effort should be put forth to preemptively minimize the possibility of a negative first review. Moreover, if the first review turns out to be negative, immediate action should be taken to reverse its negative impact.

Second, while the literature has attributed various advantages to consumer reviews, such as higher credibility, up-to-date information, and nuanced communication (Chen and Xie 2008, Ranard et al. 2016), our research identifies a significant weakness of online consumer reviews. Specifically, since the availability of consumer reviews is conditional on the adoption of the product, not all products have the same chance of receiving informative consumer reviews. For example, a product with low initial sales (which can be caused simply by an unfavorable first review) would fail to generate a viable number of reviews, and such a low review volume would further harm the product's sales. Thus, if a product receives an unfavorable initial review, consumers may lose the opportunity to learn about the product altogether, which also potentially deprives them of the chance to enjoy a high-quality product. Our data set shows that 288 out of

1155 vacuum cleaner models receive a negative first review on Amazon, which suggests that about one fourth of all models offered by this platform may fail to provide informative review information. Uncovering this important weakness of online consumer reviews is crucial, because consumer reviews are gaining greater significance in today's market, and firms are increasingly paying attention to them compared to other communication channels. Our findings regarding the availability bias of consumer reviews indicates that blindly following this trend might lead to a suboptimal mix of marketing communication efforts. By recognizing this important limitation of online consumer reviews, firms could develop an effective communication strategy that takes into account the complementarity between consumer-generated information and firm-initiated marketing activities.

Third, our research advances the literature of review dynamics. While most existing studies have focused on WOM valence dynamics (e.g., Li and Hitt 2008, Wu and Huberman 2008, Godes and Silva 2012), our research examines the dynamics of both valence and volume and, more importantly, jointly considers the evolution of both. This approach allows us to uncover the path dependence of both valence and volume from the very first review, which cannot be explained by considering the evolution of the valence independently of that of the volume. Moreover, studies of the dynamics of WOM valence have suggested opposite patterns: a downward time trend (i.e., Wu and Huberman 2008, Li and Hitt 2008) or an upward time trend (i.e., Godes and Silva 2012). Our paper, in contrast, proposes and empirically demonstrates that both a downward and an upward trend can exist simultaneously for the same product: WOM valence increases with time on a platform where the first review is negative, but decreases with time on a platform where the first review is positive. A recent paper (Moe and Schweidel 2012) finds a positive relationship between the valence of the ratings environment and review-posting incidence. By modeling the interdependence between WOM valence and volume, our research demonstrates that (a) a (dis)advantage in the first review will transfer to a (dis)advantage in both WOM valence and volume, which can explain a positive relationship between the valence of the ratings environment and review-posting incidence; and (b) the (dis)advantage in WOM volume that results from the (dis)advantage of the first rating increases rather than decreases over time. The latter finding, fundamentally driven by the joint evolution of WOM valence and volume, cannot be explained by consumer motivation (Moe and Schweidel 2012).

Finally, our results highlight the asymmetric influence of positive and negative first reviews on the information value of the online review process in general. A positive first review, regardless of how biased it might be, facilitates informative online WOM. Even with a significant upward bias, a positive first review leads to a high future WOM volume, which, as more reviews are posted, will effectively correct any initial bias. However, a negative first review is destructive to online WOM. An unfavorable first review reduces the number of consumers who will experience the product and potentially provide additional reviews. As a result, there is little chance to correct the initial bias, and the online WOM will fail to offer sufficient product information. This finding is in stark contrast with the work of Muchnik et al. (2013), who observe bias correction in the evaluation of news article comments after a negatively manipulated initial rating but positive herding after a positively manipulated initial rating. The crucial difference is that the valence and volume of the submitted evaluations in their context are independent of each other. In consuming news content or comments on news, both positive and negative ratings could equally attract readers, because the consumption of the content is initiated by their own interest and not by what others think. In the context of product reviews, however, the valence of previous reviews affects sales in the subsequent periods, because consumers tend to prefer what they have been told are better products. In addition, an evaluation can be made conditional on the purchase of the product, and thus, the valence of previous reviews affects the volume of subsequent reviews through sales. Review volume, in turn, influences the overall valence of reviews, since bias in the first review may or may not be corrected by future review volume. Therefore, interdependence between volume and valence is key to understanding the first-review effect in the context of online product reviews.

The rest of this paper is organized as follows. In Section 2, we develop a theoretical model to study the interdependence of valence and volume of online WOM, which generates predictions regarding the impact of the first review. In Section 3, we empirically test each of these theoretical predictions using data collected from the field. Finally, in Section 4, we summarize our findings and discuss future research opportunities.

2 Theoretical Analysis

To develop a conceptual understanding of the first review effect, in this section we provide a theoretical analysis of online WOM. Our model is highly stylized for the purpose of illustration, but includes essential components of online WOM evolution well-documented by extant empirical studies.

2.1 A Model of Online WOM

We first examine the evolution of online WOM. To focus on the evolution process, we consider a case in which consumers make a dichotomous purchase decision regarding a single product in the category. We first describe how previous reviews affect consumers' purchase decisions and then how reviews are updated.

Consider a consumer who evaluates a product to purchase in period t. Her valuation of the product depends on both the perceived quality and the fit of the product with her taste. First, she infers the quality of the product by reading the reviews that previous users have written (e.g., Chevalier and Mayzlin 2006, Mizerski 1982). However, because of the subjective nature of the reviews, previous reviews provide a noisy signal of quality. Thus, product quality, as perceived by a review reader, is given as $r_{t-1} + \eta_t$, where r_{t-1} is the average rating of the product in period $t-1$ and η_t is the noise in the reviews, which follows $N(0, \sigma_t^2)$. Note that the noise in the review tends to be smaller for products with extreme rather than intermediate qualities, and that it decreases as more reviews arrive (Bikhchandani et al. 1998, Zhang and Liu 2012, Zhang et al. 2015). Thus, we further model the variance of the review noise as $\sigma_t^2 = \frac{\alpha(1-\alpha)}{n_{t-1}}$, where α represents the true quality of the product and n_{t-1} is the number of reviews in period $t-1$. Note, however, that consumers only observe σ_t^2 by reading reviews but do not necessarily know the exact formulation of the variance. Second, each consumer knows her exact preference and evaluates the fit of the product based on that preference. A consumer lowers her valuation of the product in proportion to the mismatch between her own taste and the product, which we denote by x. Since consumers are heterogeneous in their taste, we assume that x follows $U[0,1]$. Together, the valuation of a consumer is given by $v_t = r_{t-1} + \eta_t - \tau \cdot x$, where τ is the mismatch cost.[3]

Based on her valuation v_t and the price p of the product, a consumer makes a purchase decision. Given some uncertainty concerning product quality, we assume that consumers are risk-

---

[3] We note that a consumer may obtain some information on the fit of the product from reading review texts (Sun 2012, Chen and Xie 2008). However, the fit information may not be available in every review text, and it is hard to identify the fit information separately from the quality information in the review ratings. Thus, we focus on the category in which quality carries much more weight in the purchase decision than fit and assume that reviews affect the purchase decision only through quality perception. However, even if the fit information is included, the first-review effects will persist, because the mechanism we uncovered for the first-review effect is the interdependence between the valence and volume of consumer reviews, which holds true irrespective of how consumers use the fit information in reviews.

averse and derive the following exponential utility: $U_t = 1 - \frac{1}{c} e^{-(v_t - p)}$, where c is a constant. Note that we adopt an exponential utility formulation to represent a risk-averse agent's utility (see for other applications of this utility formulation, Feng and Xie 2012, Holmstrom and Milgrom 1987, Holmstrom and Tirole 1993, and Kornish and Li 2010). In this specific formulation, for simplicity, we have normalized the Arrow-Pratt risk aversion parameter to be one. Later in our analysis, we also use another normalization $c = e$, in order to ensure non-negative sales for all possible review valences. Given this utility, the consumer buys the product if and only if the expected utility is non-negative: $EU_t = 1 - \frac{1}{c} e^{-\left(r_{t-1} - \tau x - p - \frac{1}{2}\sigma_t^2\right)} \geq 0$ (see Appendix A for detailed derivation).

In each period, M consumers arrive in the market and make purchase decisions as described above. However, not all consumers who make a purchase will write a review. According to a recent empirical study, only 1.5 percent of buyers write reviews (Anderson and Simester 2014). To capture the review-updating process in a parsimonious way, we use a constant propensity to write a review, $\delta \in (0,1)$, and derive the number of newly arriving reviews in period t as $\delta \cdot s_t$, where s_t represents the sales in period t.[4] Note that the valence of each review is determined by each individual consumer's experience with the product. Thus, individual-specific factors might influence the valence of an individual review. However, the reviews would reflect the true quality of the product if the entire population of buyers were to write a review. Hence, we model the valence of an individual review as a sampling process from a Bernoulli random variable with the success probability given as the true quality $\alpha \in (0,1)$. In other words, the value of each individual review is dichotomous: it takes either 0 (negative) or 1 (positive). This implies that the number of positive reviews written in period t follows the binomial distribution: $B(\delta s_t, \alpha)$. Then, the valence of the review in period t, denoted by r_t, is given as the proportion of the positive reviews in the entire set of reviews in period t. Note that this formulation ensures that consumer reviews, as a whole, are informative and furthermore, become more informative with greater volume and with product quality either higher or lower, which is consistent with our earlier assumption on the

[4] We choose to model the aggregate review-posting propensity rather than individual consumer's review-posting decision, because there have been conflicting empirical findings on the review-posting incentives (e.g., Moe and Schweidel 2012, Wu and Huberman 2008). Moreover, to date, no empirical observations have been made on any systematic variation in the propensity of writing a review across time. Even if there had been, however, our model could easily accommodate the period-specific propensity and, furthermore, our main results would not change.

review variance. However, even when consumer reviews are uninformative (i.e., a positive review is equally likely as a negative review), all our results hold (see Appendix A for proof).

To focus on demand-side dynamics, we do not model the firm's pricing decision. However, even with the endogenous price, all our results remain robust. While it is possible that some firms may adjust their prices in response to reviews, we do not find such a pattern from our data (refer to Appendix B for details of this empirical analysis). [5] Finally, to ensure the non-negativity of sales, we assume that the price is bounded between zero and $1/2$.

2.2 Evolution of Online WOM

Based on the model described thus far, we investigate the evolution of online WOM. We specifically examine how the two key variables of online WOM (valence and volume) evolve from period to period. We start with the valence and the volume of period $t-1$: r_{t-1} and n_{t-1}. First, note that these two metrics of online WOM affect the sales of period t. Since consumers purchase the product if and only if $EU_t \geq 0$, the sales of the product in period t are given by $s_t = \frac{M}{\tau}\left(1 + r_{t-1} - p - \frac{\alpha(1-\alpha)}{2n_{t-1}}\right)$ (see Appendix A for detailed derivation). Next, recall that, in every period, on average, δ proportion of buyers write reviews. Thus, the (expected) number of reviews in period t is given by

$$n_t = n_{t-1} + \delta \cdot s_t = n_{t-1} + \frac{\delta M}{\tau}\left(1 + r_{t-1} - p - \frac{\alpha(1-\alpha)}{2n_{t-1}}\right). \tag{1}$$

Similarly, since each review is positive with probability α, the expected number of positive reviews among $\delta \cdot s_t$ newly arriving reviews in period t is $\alpha \cdot \delta \cdot s_t$. Then, by denoting the total (expected) number of positive reviews in period t by h_t, we have $h_t = h_{t-1} + \alpha \cdot \delta \cdot s_t$. While, in reality, the actual number of reviews (and positive reviews) may deviate from the above-derived expected values, to capture the average effect, we integrate out randomness over a sufficiently large segment of time and use the expectations to describe the transition of the number of reviews. Finally, we simply calculate the average rating at given n_t and h_t as follows:

$$r_t = \frac{h_t}{n_t} = \frac{2n_{t-1}\{\tau h_{t-1} + \alpha\delta M(1 + r_{t-1} - p)\} - \alpha\delta M\alpha(1-\alpha)}{2n_{t-1}\{\tau n_{t-1} + \delta M(1 + r_{t-1} - p)\} - \delta M\alpha(1-\alpha)}. \tag{2}$$

Equations (1) and (2) describe the period-to-period transition of the valence and the volume,

[5] If, hypothetically, the seller adjusts its price based on the review valence, we expect the first-review effect to become smaller but it will not completely die out. This is because the higher price after the positive first review may dampen the sales increase but it is never optimal for the seller to overshoot the price such that it loses the advantage in sales obtained from the positive first review over the negative first review.

which determines the over-time evolution of these two metrics. Interestingly, these equations show that the current state of both metrics is affected by the previous state not only of their own but also of the other metric. An important implication of this observation is that the valence and the volume of online WOM evolve interdependently. On systematically investigating the relationship between the two variables, we obtain the following theorem.

Theorem 1 (Interdependence Between Valence and Volume of Online Consumer Reviews)
Two key metrics of online consumer reviews, valence and volume, evolve interdependently with positive or negative feedback. Specifically, in any given period,

(a) the current valence (r_t) positively influences the future volume (n_{t+1});

(b) the volume of newly posted reviews ($n_{t+1} - n_t$) influences the valence (r_{t+1})

  i. positively if the previous valence is biased downwards (i.e., $r_t < \alpha$),

  ii. negatively if the previous valence is biased upwards (i.e., $r_t > \alpha$).

For proofs, see Appendix A. The theorem describes how the two metrics influence each other in their evolution. To understand Part (a) of the theorem, recall from the extant research that positive reviews from the previous period enhance the product's sales (Babić et al. 2016, You et al. 2015): $\frac{\partial s_{t+1}}{\partial r_t} > 0$. Moreover, note that only consumers who have purchased the product post online reviews, implying that each review posting is conditional on the adoption of the product. Therefore, the volume of consumer reviews in any given period will be positively affected by the level of sales: $\frac{\partial n_{t+1}}{\partial s_{t+1}} > 0$. [6] This finding implies that online consumer reviews may provide insufficient information, depending on the current level of sales. Especially, in the case of low sales, consumers could suffer from this information availability bias. Together, these two inequalities imply that the valence of consumer reviews increases the volume of reviews in the subsequent period: $\frac{\partial n_{t+1}}{\partial r_t} = \frac{\partial s_{t+1}}{\partial r_t}\frac{\partial n_{t+1}}{\partial s_{t+1}} > 0$.

Part (b) of the theorem discusses the impact of the volume on the future valence of the online reviews. In this case, unlike in Part (a), the impact may be either positive or negative, depending on the magnitude of the current valence relative to the true quality of the product. To understand this, note that, in any period, the valence is the weighted average of the valence of all previously posted reviews and that of newly posted reviews, with the weight being the volume of each group of reviews. Thus, if the valence in the previous period is more negative than that of the

[6] We acknowledge that some reviews might be written without purchase (Anderson and Simester 2014). However, the presence of reviews without purchase does not change the impact of the current valence on the future volume.

newly posted reviews, a greater volume of newly posted reviews will render the valence of the current period more positive. However, if the valence in the previous period is more positive than that of the newly posted reviews, a greater volume of newly posted reviews will turn the valence of the current period more negative. On average, the valence of the newly posted reviews is consistent with the true quality of the product. Thus, if the realized valence of already posted reviews happens to be lower than the product's true quality, the newly posted review volume will elevate the future valence. However, if the realized valence is higher than the true quality, the newly posted review volume will lower the future valence.

An important implication of Theorem 1 is that one cannot separately trace the evolution of the two key metrics of online WOM, valence and volume. Rather, their evolution patterns can be accurately captured only when their interaction is considered. Given the evolution patterns characterized thus far, it is evident that previous reviews can influence both the valence and the volume of subsequent reviews. Such influence leads to the possibility that the very first review might have a significant impact on the evolution of both the valence and the volume of the entire set of reviews. We examine this possibility in the next subsection.

2.3 The First-Review Effect

In this section, we examine how the first review might influence both the valence and the volume of subsequent reviews. For this purpose, we start from the initial period ($t = 0$), where the very first review is posted. After this initial period, the first review is exogenously given as either positive ($r_0 = 1$) or negative ($r_0 = 0$). Thus, depending on the valence of the first review, we separately trace the evolution of valence and volume using superscript $k \in \{+, -\}$, where $k = +$ if $r_0 = 1$ and $k = -$ if $r_0 = 0$. Then, based on equations (1) and (2), we can derive the general expression of the series n_t^k and r_t^k as follows: $n_t^k = 1 + \delta \sum_{i=1}^{t} s_i^k$ and $r_t^k = \frac{h_0^k + \alpha \cdot \delta \sum_{i=1}^{t} s_i^k}{1 + \delta \sum_{i=1}^{t} s_i^k}$, where $s_t^k = \frac{M}{\tau}\left(1 - p + \frac{2 \cdot \left(h_0^k + \alpha\delta \sum_{i=1}^{t-1} s_i^k\right) - \alpha(1-\alpha)}{2 \cdot \left(1 + \delta \sum_{i=1}^{t-1} s_i^k\right)}\right)$. By comparing n_t^+ and n_t^- as well as r_t^+ and r_t^-, we obtain the following finding on the impact of the first review on the valence and the volume of online WOM in every subsequent period.

Proposition 1 (First-Review Effect: Overall Impact)

A product's first consumer review persistently influences its future consumer reviews. Specifically, a higher rating in the first reviews leads to (a) a higher average rating and (b) a higher number of reviews in any given period. Formally,

$$r_t^+ - r_t^- > 0 \text{ and } n_t^+ - n_t^- > 0, \forall t \tag{3}$$

The above result shows that both the valence and the volume of WOM crucially depend on the valence of the first review. More specifically, when the first review is positive, both the average rating and the total number of reviews are higher than when the first review is negative. Below, we provide intuitions for the valence and the volume results, respectively.

In general, many idiosyncratic factors affect the valence of each individual review. Thus, we do not expect the first review to affect the average rating in subsequent periods in any significant way. Nevertheless, the first part of the proposition states that a positive first review generates a higher average rating in every subsequent period. To understand the rationale, recall that the valence in any period is the weighted average of the valence of all previously posted reviews and the average rating of newly posted reviews. Also, note that each individual review is independent of previous reviews and, thus, the average rating of newly posted reviews does not depend on the valence of the first review. Therefore, any advantage in the average rating of the previous period is likely to remain in the average rating of the next period, implying that the average rating in subsequent periods will be more positive if the first review is positive than if it is negative.

Proposition 1 also shows that a greater number of reviews will be generated after a positive than after a negative first review. Two different mechanisms contribute to this result. First, recall from the first part of the proposition that a higher rating in the first review results in higher average ratings in subsequent periods. Since consumers are more likely to purchase a product when online reviews are favorable (Babić et al. 2016, You et al. 2015), sales will be greater after a positive than after a negative first review. Then, in the subsequent period, a greater volume of WOM will follow. Second, more reviews reduce consumers' uncertainty regarding the quality of the product, regardless of the valence of the reviews, because more reviews reflect greater sales, which in turn work as a social proof that the product quality is sufficiently high to warrant a purchase (Zhang and Liu 2012). Thus, consumers are more attracted to a product when making a purchase decision if they see a large number of posted reviews. Hence, more reviews lead to greater sales and, in turn, to a greater volume of WOM, which establishes a positive feedback loop between sales and WOM volume. Both of these two effects (the direct impact of WOM valence on sales and the positive feedback loop between sales and volume) work together to generate a greater volume of WOM from a positive than from a negative first review.

Given the well-documented impact of WOM valence and volume on sales (Babić et al.

2016, You et al. 2015), these results have important implications. Our results suggest that a product of the same quality may generate different amounts of WOM with different average ratings, thus achieving a different level of sales, all depending on the valence of the first review. More importantly, this impact persists over time; thus, any negative impact due to a first negative review is very hard to reverse without costly intervention. Therefore, when launching a new product, it is critical that the first review be positive.

Given that the first-review effect persists, an interesting question is how the magnitude of this effect changes over time. To examine this issue, we study the evolution of the valence and the volume of consumer reviews and obtain the following result.

Proposition 2 (First-Review Effect: Over-time Dynamics)

(a) The first-review effect on valence diminishes over time:

$$\Delta r_t > \Delta r_{t+1}, \forall t, \tag{4}$$

where $\Delta r_t = r_t^+ - r_t^-$ and $\Delta r_{t+1} = r_{t+1}^+ - r_{t+1}^-$.

(b) The first-review effect on volume intensifies over time:

$$\Delta n_t < \Delta n_{t+1}, \forall t, \tag{5}$$

where $\Delta n_t = n_t^+ - n_t^-$ and $\Delta n_{t+1} = n_{t+1}^+ - n_{t+1}^-$.

The above result characterizes the evolution of the valence and the volume of WOM as a function of the valence of the first review. The first part of the proposition shows that the impact of the first review on the valence of WOM becomes weaker as time passes. This weakening occurs because any bias in the first review is corrected as more reviews arrive. Recall that the valence of an individual review is drawn from a Bernoulli distribution with a success probability α (i.e., the product's true quality). The law of large numbers suggests that the average rating will converge to this true quality over the long term. Thus, if the average rating is too high at the beginning ($r_0 = 1$) it will decrease to α; but if it was initially too low ($r_0 = 0$), it will rise to α. Therefore, the first-review effect on the valence becomes weaker over time, although, as Proposition 1 suggests, it does not completely disappear.

In contrast, the first-review effect on WOM volume not only persists, but is also reinforced as time passes, according to the second part of the proposition. This result is obtained because, in every period, the number of newly arriving reviews is greater when the first review is positive than when it is negative. Note that Proposition 1 discusses the *total* volume of WOM in each period as a function of the first review. However, the same intuition applies to the newly arriving reviews in each period. In particular, since by Part (a) of Proposition 1, the positive first review shifts the valence of overall WOM upwards, consumers are more likely to purchase the product and, thus, the number of newly arriving reviews in every period is likely to be greater when the first review is positive than when it is negative. Therefore, due to the difference in the first review, the incremental number of reviews accumulates in each period, which strengthens the impact of the first review on the overall volume period by period. This result implies that a small difference at the beginning could lead to a drastic divergence at the end. In this sense, the online WOM phenomenon is subject to a severe path dependence, where the very first review determines the evolution path, and thus the fate of a product.

Overall, our theoretical analysis implies that a negative first review is detrimental to the performance of a product in several different ways. First, with a negative first review, the valence of WOM remains negative for a significant amount of time, which reduces sales, according to recent meta-analyses (Babić et al. 2016 and You et al. 2015). At the same time, the volume of WOM for the product is lower, which again leads to lower sales, as shown by the aforementioned meta-analyses. Moreover, even though a negative bias in the average rating of a product can be corrected over time, it may have little or no chance of correction because of a lower review volume. In an extreme case, a product may not even take off as a result of a negative first review, resulting in only mediocre sales at best, and thus may not obtain many additional reviews, keeping the online sentiment towards the product negative. Therefore, we should not consider the first review as just a *single* review, but rather as *the most influential* review that, as such, must be properly managed.

2.4 Informativeness of the First Review

So far, we have examined the first-review effect by comparing the evolution of the online WOM following the two possible realizations of the first review (i.e., positive and negative). This formulation allows us to remain agnostic about whether the first review is informative of product quality. This suggests that our results on the first-review effect do not depend on any specific assumption regarding the informativeness of the first review. However, since the informativeness of the first review can also vary as a result of environmental factors, it is useful to examine its implications by considering the following two types of the first review: informative vis-a-vis uninformative.[7]

The first review is considered informative when its valence reflects the true quality of the product (i.e., α). In this case, the first review is positive with probability α but negative with

[7] We thank the review team to motivate us to study this issue.

probability $1 - \alpha$. In contrast, when the two realizations of the first review are equally likely, the first review is uninformative. Since, according to the first-review effect, a positive first review is likely to generate more sales than is a negative first review for an identical product, the expected sales would differ depending on the informativesness of the first review. In particular, a high-quality product (with $\alpha > 0.5$) will have greater sales when the first review is informative than when it is uninformative. On the contrary, a low-quality product (with $\alpha < 0.5$) will achieve lower sales when the first review is informative than when it is not (see Appendix A for formal proofs of these claims). This result implies that a high-quality firm would prefer an informative first review while a low-quality firm would prefer an uninformative one.

Therefore, in order to promote high-quality products, an e-commerce platform may encourage early consumers to write a more informative first review. However, in reality, the first review often turns out to be random. As shown in the introduction, in many cases, first reviews for identical products are inconsistent across different platforms. Also in the data set we use for our empirical analysis, we find no significant correlation between the first review and the quality of the product (as measured by the Consumer Reports ratings).[8] These findings underscore the importance of understanding the first-review effect, because it has direct implications to firms and consumers.

So far, our theoretical analysis has generated four testable predictions regarding the significance of the first review, which we present in Table 2 (A1~A4). Given their counterintuitive nature, it is necessary to subject each of these predictions to empirical tests. In the next section, we present empirical studies to test these predictions.

3 Empirical Analysis

In this section, we report results of our empirical investigations into the first-review effect. We use vacuum cleaners as our empirical context in our analysis because, in this category, consumers' purchase decisions are likely to be heavily influenced by product reviews. Note that vacuum cleaners are privately consumed, durable goods with infrequent product trials and thus, are susceptible to online WOM (You et al. 2015). Furthermore, our analysis of review texts reveals that consumers of vacuum cleaners consider more vertical than horizontal attributes in their

---

[8] However, the correlation between the average rating (across all reviews) and the quality of the product is significant. The no-correlation result (of the first review) is likely due to a small sample size.

purchase decisions (see Appendix B for details), which indicates that the vacuum cleaner category provides an appropriate context to test our theory.

To identify the first-review effect, we compare the volume and the valence of WOM when the first review is positive with those when it is negative. Our study specifically uses a sample of products that simultaneously appear on two different retail platforms, Amazon and Best Buy, and examines how different first reviews on different platforms could make a difference in the valence and the volume of the online reviews for the same product. Since we compare the WOM metrics for the same product, we can identify the first-review effect while perfectly controlling for any unobserved heterogeneity specific to the product (see also Chevalier and Mayzlin 2006 for a similar approach). The platform-specific effects are also differenced out in our estimation. In what follows, we first present our data as well as initial evidence that our theoretical predictions are in line with the empirical observations. We then describe formal tests of the theory in greater detail. Later, we also discuss the robustness of our results.

3.1 Data

We collect online review information as well as product characteristics on vacuum cleaners from both Amazon and Best Buy. For completeness, we cover in our data set all the vacuum cleaners with at least one review from both Amazon and Best Buy at the time of data collection (January 6, 2015). This results in a total of 177 vacuum cleaner models in our data set. For each of these products, we collect the number of reviews, the order of each review, the date each review was posted, the valence of each review, the review texts, and the price of the product from both platforms. Table 3 provides summary statistics of this data set.

Given this data, for each product, we calculate the average rating as well as the number of reviews each month since the first review was posted, which allows us to have multiple observations of the average rating and the number of reviews for each product across time. Note, however, that different products in our data set have a different number of observations because they were launched, and their first reviews were posted, at a different point in time. Thus, for products launched closer to the data collection time, observations are truncated earlier, thus resulting in a smaller number of observations. While we do not expect this data structure to cause any selection bias in the overall impact of the first review, the evolution pattern of WOM might be heavily influenced by products with longer durations. To minimize such an asymmetric influence while maintaining the representativeness of the sample, we confine our observation window to one year, during which about 71 percent of the products in the sample cover the whole observation period, while the remaining 29 percent have observations truncated at some point before the end of the observation period. In Section 3.5.1, we consider observation windows of different lengths to show the robustness of our findings.

3.2 Initial Evidence

Our theory concerns the valence and the volume of WOM as a function of the valence of the first review. We classify the first review as positive if its rating is strictly greater than three stars, but as negative otherwise, following Amazon's dichotomization scheme. In our data set, 31.07 percent of 177 products exhibit inconsistent first reviews across the two platforms, according to this classification. Specifically, 21 products obtain a positive first review on Amazon and a negative one on Best Buy, while 34 products have a negative first review on Amazon and a positive one on Best Buy. By examining these two groups of products, one can potentially identify the first-review effect. We thus calculate both the average star rating and the average number of reviews on each platform for these two groups. We report the results in Table 4. First, Table 4 shows that the first group of products (with a positive first review on Amazon and a negative one on Best Buy) has a higher average rating as well as a higher number of reviews on Amazon than on Best Buy. This observation is in line with the first-review effect (i.e., a positive first review leads to more reviews with a higher average rating). Second, the other group (with a negative first review on Amazon and a positive one on Best Buy) exhibits a lower average rating but a higher number of reviews on Amazon than on Best Buy. While the average rating moves in the same direction as our theory, the result on the number of reviews seems to contradict the first-review effect. However, Table 4 also suggests that this might be due to a platform-specific effect. Specifically, we find that the products sold from both platforms on average receive 0.26 fewer stars but 48.49 more reviews on Amazon, across all samples. Even if the first reviews are consistent across the two platforms, Amazon induces more reviews, while Best Buy generally generates higher ratings (see the third and the fourth rows of Table 4), which implies that a product may receive a lower average rating and a higher number of reviews simply by being posted on Amazon rather than on Best Buy.[9] Thus, for a fair comparison, the platform-specific effect needs to be controlled. To account for this, we additionally conduct a non-parametric test that does not depend on the platform-specific effect.

[9] While the fourth column of Table 4 suggests the (non-significant) opposite direction for the comparison of the average rating, this is likely due to a relatively small sample size.

We will also formally address this issue when we develop our empirical model in the next section.

As an alternative test, we conduct the Wilcoxon signed-rank test. In this test, we can avoid the problem of unequal baselines across the two platforms by using ranks instead of raw data. Thus, for the two groups of products with different first reviews across the two platforms, we test whether or not the average ranks of the products on one platform are comparable to those of the same products on the other. However, since both the average rating and the number of reviews vary over time, to determine the ranking of products, we fix the time as the end of one year and consider only products that reach this particular time period. In this sample, 37 products have inconsistent first reviews. The first row of Table 5 reports the results from the comparison of these 37 products, where lower numbers imply higher rankings. As shown in this table, products typically rank higher for both the average rating and the number of reviews across the two platforms if the first review is positive rather than negative. To see whether these results are specific to the observation timing or the size of the sample, we also run the same test for a sample of 46 products that reached the end of the sixth month, and report the results in the second row of Table 5. As can be seen from the table, we obtain qualitatively the same results, i.e., in both metrics, the rankings are significantly higher for the same product when its first review is positive, rather than negative. Therefore, our preliminary analysis of the data supports the first-review effect: A positive first review generates more reviews with higher ratings than does a negative first review.

3.3 The First-Review Effect on the Valence of WOM

3.3.1 Model and Estimation

In this section, we formally test our predictions on the impact of the first review on the valence of WOM (i.e., A1 and A3). Recall that A1 states that the average rating of a product is higher when its first review is positive than when it is negative. The same hypothesis also implies that this should hold for every period. Moreover, A3 states that the first-review effect on the average rating diminishes over time. To test the validity of the predictions in these hypotheses, we model the average rating as a function of (a) the valence of the first review, (b) the time passed since the first review was posted, and (c) their interaction. Note that the average rating of a product changes with time, as new reviews arrive. Thus, the average rating is specific not only to the product, but also to the elapsed time since the first review was posted. Therefore, we adopt the panel structure and consider the following model:

$$AR_{it}^{j} = \beta_0^{j} + \beta_1 FNegative_i^{j} + \beta_2^{j} \ln(FDuration_{it}) + \beta_3 FNegative_i^{j} \times \ln(FDuration_{it})$$

$$+X_{it}^{j}\Gamma^{j} + \varepsilon_{it}^{j} \quad (6)$$

where subscript i is the indicator for the product, subscript t is the indicator for the period, and superscript j is the indicator for the platform ($j = \mathrm{A}, \mathrm{B}$; A is Amazon and B is Best Buy). $FNegative_{i}^{j}$ is a dummy variable representing the valence of the first review of product i in platform j, following Amazon's dichotomization scheme presented above (1=negative; 0=positive). $FDuration_{it}$ represents the duration of the first review of product i at period t, as measured by the number of months. Note that this is not a calendar date but the elapsed time since the first review posting. Moreover, $FDuration_{it}$ always takes the same value as the period: $FDuration_{it} = t$. In the regression, due to a curvilinear relationship between the average rating and the duration of the first review, we use the log transformation of the duration.[10]

The matrix X_{it}^{j} contains control variables, including the log of the product price, the log of the number of words in the first review (WC), consumer search volume on Google (both at product and brand levels), product type dummies, product feature indicators, and the volume proxy.[11] We include these variables to control product-specific and product- and platform-specific effects that might affect the evolution of online WOM. For example, a product's price might correlate with the valence of its reviews; the first review may be more influential if its review text is lengthy; consumer search volume might correlate with underlying demand for the product and thus influence the volume of WOM; and it is also possible that certain product types and features could influence the popularity of the product and thus, the volume of WOM. Finally, we use the volume proxy to control the impact of the WOM volume on the average rating. However, using the volume itself will distort the estimate for the first-review effect on the valence, since the effect through the volume will be cut out (see Chapter 9 of Gelman and Hill 2006, for more details). Hence, we chose to use, as a proxy, the residual from the regression of the volume on $FNegative$, $\ln(FDuration)$, and all controls in the matrix X_{it}^{j} other than the volume proxy. Note that, to keep consistency with our theoretical model, we have used the lagged residual as the volume proxy.

---

[10] Even without the log transformation, we could obtain qualitatively same results, although the current model with the log transformation fits better with the data.

[11] Among these controls, the product type dummies are indicator variables for vacuum cleaner types as classified by Amazon: Canister, Handheld, Robotic, Stick, or Upright. For identification, we include only four indicators, while setting Canister as the baseline type. As the product feature indicators, we include two dummies: one for bagged vacuum cleaners and another for corded vacuum cleaners. These are coded as one if the product has the feature but zero otherwise.

Note that, in our empirical model, platform-specific effects are captured by the intercept (β_0^j) and the coefficient for duration term (β_2^j), with β_0^j representing the time-invariant baseline rating for the platform and β_2^j capturing the time-varying part of the platform-specific effect. Currently, none of these platform-specific effects are captured by the first-review effect coefficients (i.e., β_1 and β_3), but we later allow these first-review effect coefficients to be platform-specific as well (see Section 3.5.7). Lastly, the errors of the model (ε_{it}^j) represent an unobserved heterogeneity specific to both the product and the platform that is not explained by the independent variables. We assume the errors to be normally distributed with mean zero.

As previously noted, in this study, the identification of the first-review effect is based on the variation across websites for the same product. To implement this identification strategy, we eliminate product-level fixed effects by differencing the average ratings in the two platforms for each product i and each period t, which leads to the following estimation equation:

$$\Delta AR_{it} = \beta_0 + \beta_1 \Delta FNegative_i + \beta_2\, ln(FDuration_{it}) + \beta_3 \Delta FNegative_i \times ln(FDuration_{it})$$
$$+ \mathrm{X}_{\mathrm{it}}^{\mathrm{A}}\Gamma^{\mathrm{A}} - \mathrm{X}_{\mathrm{it}}^{\mathrm{B}}\Gamma^{\mathrm{B}} + \varepsilon_{it} \quad (7)$$

where $\beta_0 \equiv \beta_0^A - \beta_0^B$, $\beta_2 \equiv \beta_2^A - \beta_2^B$, $\Delta AR_{it} \equiv AR_{it}^A - AR_{it}^B$ and $\Delta FNegative_{it} \equiv FNegative_{it}^A - FNegative_{it}^B$. Note that, for the control variables taking the same value across the two platforms (i.e., product type dummies and product feature indicators), Γ^{A} and Γ^{B} will not be separately identified, but only their difference will be. For other controls, both Γ^{A} and Γ^{B} can be separately identified. The error term ($\varepsilon_{it} = \varepsilon_{it}^A - \varepsilon_{it}^B$) is the difference in the error terms and, thus, is also normally distributed with mean zero. Therefore, the error term still captures unobserved heterogeneity specific to the product. This also implies that errors are likely to be correlated across periods within a product. In this case, OLS still produces consistent estimates of the model parameters, but the standard errors will be underestimated (Bertrand et al. 2004). Hence, we estimate our main parameters using OLS, but we also cluster standard errors by product to account for potential serial correlation (Liang and Zeger 1986). In our estimation, we only use 138 products from our data set that receive at least two reviews from both platforms, since with only one review, the average rating and FNegative are perfectly correlated. Note that these 138 products include those with inconsistent first reviews across the two platforms, and also those with consistent first reviews, with the former identifying the first-review effect and the latter capturing the baseline difference across the two platforms.

3.3.2 Results

We report our estimation results in Table 6. The table shows that the data support both of our predictions regarding the valence of WOM (i.e., A1 and A3). First, the coefficient for $\Delta FNegative$ (β_1) captures the impact of the difference in the first review on the average rating. Given that $FNegative_i^j$ takes a value of one when the first review is negative in platform j but zero otherwise, and we expect β_1 to be negative according to A1. Table 6 shows that the estimated value of β_1 is negative and statistically significant ($\beta_1 = -1.432, p < 0.01$), implying that a product receives a more unfavorable review on a platform where its first review is negative than on another platform where its first review is positive. This result confirms the first-review effect on the valence of WOM (i.e., A1).

Second, we can test the first-review effect on the over-time dynamics of the valence (i.e., A3) by the coefficient for the interaction of $\Delta FNegative$ and $\ln(FDuration)$ (i.e., β_3). According to A3, the first-review effect on the valence decreases over time and, thus, β_3 is expected to be in the opposite direction of the main effect (i.e., β_1). Since β_1 is negative, we expect β_3 to be positive. Consistent with this prediction, the coefficient for the interaction (β_3) is positive and significant ($\beta_3 = 0.498, p < 0.01$). This result indicates that the difference caused by the first review decreases over time. Yet, the effect does not disappear for an extended period of time, as A1 suggests. For example, our estimation results predict that receiving a negative (rather than positive) first review for a product reduces the average rating by 0.54 stars by the end of the sixth month and by 0.19 stars by the end of the first year. According to our estimates, even after 18 months, the average rating is expected to be lower when the first review is negative than when it is positive. Thus, the first-review effect persists over time. Even though the average rating should theoretically converge to the true quality in the long run, it would take a significant amount of time to do so, because a negative first review induces a lower volume of subsequent reviews, as shown in Proposition 1(b). Therefore, the valence of the first review has a non-trivial and persistent impact on the valence of the overall WOM.

Finally, recall that the intercept (β_0) and the coefficient for $\ln(FDuration)$ (β_2) together capture the difference in the platform-specific fixed effects. According to Table 6, the estimate for β_0 is not significant but that for β_2 is negative and significant ($\beta_2 = -0.130, p < 0.01$). This result suggests that the average rating of a product is higher on Best Buy than on Amazon in every period after the initial period and that this gap increases with time, regardless of the valence of the first review. More importantly, our result shows that the first review has a significant and long-lasting impact on the valence of WOM, even after controlling for platform-specific effects.

3.4 The First-Review Effect on the Volume of WOM

3.4.1 Model and Estimation

We now turn to the impact of the first review on the volume of WOM. In particular, we test A2 and A4: in any period, the number of reviews is greater when the first review is positive than when it is negative; the gap in the number of reviews not only persists, but also intensifies as time passes. To test the validity of these predictions, as in the valence study, we model the total number of reviews received up to period t as a function of the duration of the first review and its interaction with the valence of the first review along with controls. In particular, we consider the following model: for platform j, $(j = A, B)$,

$$\ln\left(NR_{it}^{j}\right) = \beta_0^j + \beta_1 FNegative_i^j + \beta_1^j FDuration_{it} + \beta_2 FNegative_i^j \times FDuration_{it} + \mathrm{X}_{\mathrm{it}}^{\mathrm{j}}\Gamma^{\mathrm{j}} + \varepsilon_{it}^{j}, \quad (8)$$

where all independent variables as well as the matrix X_{it}^{j} containing control variables are defined as in the valence study, except that we substitute the volume proxy in X_{it}^{j} for the valence proxy. The valence proxy is constructed in the same way as the volume proxy: it is the (lagged) residual of the regression of the average rating on $FNegative$, $FDuration$, as well as all controls in the matrix X_{it}^{j} other than the valence proxy. In our model, we use the log transformation of the dependent variable since its distribution is right-skewed. The log transformation of the number of reviews variable is also standard in the WOM literature (e.g., Chen et al. 2011, Chevalier and Mayzlin 2006, Liu 2006). Note that the snowball effect is less likely to hold with the log transformation and, thus, this choice makes a more conservative test of our theory.

As in the previous study, we identify the first-review effect by differencing equation (8) of the two platforms for each product in each time period. Thus, we estimate the following equation:

$$\Delta\ln\left(NR_{it}\right) = \beta_0 + \beta_1 \Delta FNegative_i + \beta_2 FDuration_{it} + \beta_3 \Delta FNegative_i \times FDuration_{it} + \mathrm{X}_{\mathrm{it}}^{\mathrm{A}}\Gamma^{\mathrm{A}} - \mathrm{X}_{\mathrm{it}}^{\mathrm{B}}\Gamma^{\mathrm{B}} + \varepsilon_{it}, \quad (9)$$

where $\Delta\ln\left(NR_{it}\right) \equiv \ln\left(NR_{it}^{A}\right) - \ln\left(NR_{it}^{B}\right)$, and β_0, β_2, and $\Delta FNegative_i$ are defined as before. We estimate equation (9) using OLS with standard errors clustered at the product level. Unlike before, however, we suppress β_1 to be zero. This is because β_1 captures the difference in the number of reviews in the initial period due to the difference in the first-review valence, which is always zero since the volume of the first review is one irrespective of its valence. Thus, we estimate only the other parameters.[12]

3.4.2 Results

We report the estimation results in Table 7. The estimates suggest that our data support both A2 and A4. To see this, first note from equation (9), that given $\beta_1 = 0$, a product receives a smaller number of reviews in each period after a negative than a positive first review, if and only if $\beta_3 < 0$. In addition, given that β_3 is the coefficient for the interaction between $\Delta FNegative$ and $FDuration$, $\beta_3 < 0$ also implies that this gap in the volume due to the first review difference increases over time. Hence, if $\beta_3 < 0$, both the overall effect (A2) and the snowball effect (A4) will be supported. According to Table 7, the estimate of β_3 is negative and statistically significant ($\beta_3 = -0.047, p < 0.01$), thus supporting both A2 and A4. To put this result in context, our estimates specifically suggest that, if the first review is negative rather than positive, a product receives 24.6 percent fewer reviews at the end of the sixth month ($= 1 - \exp(-0.047 \times 6)$) and 43.1 percent fewer reviews at the end of the first year ($= 1 - \exp -0.047 \times 12)$). On Amazon, these numbers translate into 10.72 fewer reviews by the end of the sixth month and 43.93 fewer reviews by the end of the first year, which numbers illustrate the snowball effect.[13]

Finally, our results show that both the intercept and the coefficient of $FDuration$ are positive and significant ($\beta_0 = 2.921, p < 0.01;\ \beta_2 = 0.071, p < 0.01$). These results suggest that a significant platform-specific effect exists that affects the number of reviews. In particular, Amazon generates more reviews than does Best Buy for the same product, and the difference in the number increases over time. Despite this strong platform-specific effect, our analysis shows both overall and long-term effects of the first review on the volume of WOM.

3.5 Robustness Checks

So far, we have identified the strong influence that the first review exerts on the valence and volume of a product's entire WOM. In this section, we examine the robustness of these findings by considering different observation windows, alternative explanations (unobserved product characteristics and the display effect), other contexts (a different category and a different

[12] Even if we allow β_1 to take non-zero value, the model fit does not improve. Thus, we choose to use the model that is parsimonious but theoretically sound.

[13] However, the snowball effect may not persist towards the end of the product life cycle, which would impose the boundary condition for our theory. We observe the persistence of the snowball effect in our data set, since vacuum cleaners have a life cycle that is generally longer than the span of the data set.

pair of platforms), and alternative assumptions (temporal price variation and a platform-specific first-review effect). Our results show that the first-review effects remain robust to all these considerations and, thus, are truly robust findings. In what follows, we discuss the motivation, the method, and results for each of these analyses.

3.5.1 Observation Windows

In our main analysis, we use an unbalanced panel data set with a one-year observation window. We choose one year to balance between two objectives: minimizing heterogeneity in duration and minimizing the loss of information. However, it is important to determine how sensitive our results are to this specific choice of observation window. We thus consider two other samples with observation windows of two years and three years, and estimate Equations (7) and (9). Note that the percentage of products reaching the end of the observation period decreases, whereas the total number of reviews used in the estimation increases as the observation period becomes longer. We report the estimation results in Column 1 (two-year window) and Column 2 (three-year window) of Tables 8 and 9. In the valence study, we find that the estimates for the main effect of $FNegative$ and its interaction with $\ln(FDuration)$ are consistent in both direction and significance with those of our main analysis. In the volume study, as before, we impose the restriction that $\beta_1 = 0$ (as in all of our robustness checks). Then, the coefficient for the interaction term $\Delta FNegative \times FDuration$ is negative and marginally significant for the two-year window ($p = 0.052$) and the three-year window ($p = 0.073$).[14] Therefore, our results are robust to the length of the observation window. More importantly, by examining the first-review effects with longer observation windows, our analyses provide even stronger evidence for persistence as well as for the snowball effect.

3.5.2 Unobserved Product Characteristics

The first-review effect characterizes the relationship between the valence of the first review and the volume/valence of the entire review. However, it is possible that this relationship could be driven by differences in taste concerning certain unobserved product characteristics between two distinct consumer populations on the two platforms. If such unobserved product characteristics are indeed drivers of the relationship, the first-review effect will merely be a spurious correlation. To

[14] One possible reason for the marginal significance is the sample size. Specifically, only 25% of products reach the end of the third year, and only 35% reach the end of the second year. Thus, we estimate the snowball effect on a very limited number of products.

examine this possibility, we first note that such unobserved product- and platform-specific effects, if any, would change our results only to the extent that these unobserved effects correlate with the independent variables. Thus, the size of the potential bias would depend on to what extent variations in our covariates are attributable to these unobserved effects. Therefore, to examine the potential selection of our main variable (i.e., FNegative) on unobservables, we estimate Equations (7) and (9) excluding all control variables and then compare the estimates with those from the model with controls (i.e., our original estimates in Tables 6 and 7). Note that this no-control specification has been widely used to reduce concerns regarding selection on unobservables (e.g., Altonji et al. 2005, Cameron and Taber 2004, Mayzlin et al. 2014). We report the estimation results from the no-control specification in Column 3 of Tables 8 and 9. When we compare these results with those in Tables 6 and 7, we observe that the coefficients for the first review remain quite stable with or without inclusion of covariates. Hence, our finding on the first-review effect is unlikely to be driven by selection on unobserved product characteristics.

3.5.3 Display Effect

Recall that our theoretical model suggests that the first review makes a difference in WOM volume because a positive first review induces a higher average rating, which in turn leads to increased sales. However, it is possible that higher sales and more WOM in the case of a positive first review might stem from a greater likelihood that the product receives a first-page listing in search results on the retail platform. To examine whether or not the first-review effect survives this display effect, we estimate the model of volume with an additional control variable, the first-page dummy, to which we assign the value of one if the product appears on the first page, and zero otherwise.[15] We report the results in Column 4 of Table 9. From this analysis, we first find that the first-page display effect indeed has a significant impact on the number of reviews in Best Buy (although not in Amazon). More importantly, however, even after controlling for the display effect, we still find a significant first-review effect on WOM volume. Therefore, the first-review effect on WOM volume is robust to the presence of the display effect.

3.5.4 Different Product Category

The empirical context for our main analysis is the market for vacuum cleaners. To

[15] We use product type as a search criterion because this is the first filter that the retailers provide saliently on both Amazon and Best Buy. Moreover, the small assortment size of Best Buy makes it unnecessary to use multiple search criteria.

strengthen the external validity of our results, we also examine the first-review effect using data from a different product category. We specifically choose toasters, because toasters and vacuum cleaners share characteristics typical of a product category that is sensitive to WOM, such as private consumption, durability, and infrequent trials (You et al. 2015). As in our main study, we collect reviews and product information for all available toasters with at least one review on both Amazon and Best Buy, as of December 22, 2016. This results in a total of 113 products, among which 40.7 percent received inconsistent first reviews across the two platforms. We report summary statistics of this sample in Appendix B.

By using the toaster data, we estimate Equations (7) and (9) with observation windows of one, two, and three years. We use the same set of control variables as in the vacuum cleaner study, except for the product feature and the brand search volume on Google (which is too small compared to the category-level search and thus omitted). As the product type, we use the toaster dummy, which separates out the toaster type (1) from the oven type (0). We report the estimation results in the first three columns of Table 10 (valence) and Table 11 (volume). Table 10 shows that, in the valence model, across all three observation windows, the coefficient for $\Delta FNegative$ is negative and significant, while that for the interaction with $FDuration$ is positive and significant. Moreover, according to Table 11, the interaction between $\Delta FNegative$ and $FDuration$ in the volume model is negative and at least marginally significant ($p = 0.083$ for the one-year window; $p < 0.05$ and $p < 0.01$ for the two- and three-year windows, respectively). These results suggest that all of our predictions in A1~A4 are supported even in the toaster category. Therefore, our findings concerning the first-review effect are robust to choice of product category.

3.5.5 Different Platforms

In our main analysis, we compare reviews of the same products sold on both Amazon and Best Buy. To examine whether our findings are specific to this choice of platforms or they can be generalized, we repeat our analysis on a different pair of platforms: Amazon US and Amazon Canada.[16] By using this pair, we can also reduce the concern for unobserved product- and platform-specific effects (as discussed in Section 3.5.2). This is because the taste difference between their

[16] Amazon US and Amazon Canada have totally independent websites with separate consumer reviews even for identical products sold on both platforms. In January 2017, however, Amazon Canada started providing their customers with a link to Amazon US product reviews. Hence, in our analysis, we only use product reviews posted before January 1, 2017.

consumer populations, if such exists, is of a different type than that of the Amazon-Best Buy pairing and thus, if the first-review effect is also confirmed in this case, it is unlikely to be driven by unobserved product- and platform-specific effects. In this study, we also use the vacuum cleaner category and collect both reviews and product information available as of February 8, 2017. As before, we include in our sample all products with at least one review on both platforms, resulting in a total of 174 products. Of these, 29.9 percent received a first review of differing valence on each of the two platforms. We report summary statistics in Appendix B.

Based on this data, we estimate Equations (7) and (9) using the same set of controls as in the Amazon-Best Buy study. Note that we convert the prices in Amazon Canada into US dollars based on contemporaneous exchange rates. We also use one year as our observation window, since only a limited number of products reach the end of the second year. We report the estimation results in the first column of Table 12 (valence) and Table 13 (volume). The tables indicate that all the signs of the relevant estimates are consistent with our theoretical predictions: in the valence study, $\beta_1 = -1.736$ ($p < 0.01$) and $\beta_3 = 0.383$ ($p < 0.01$); in the volume study, $\beta_3 = -0.053$ ($p < 0.01$). Therefore, all of our predictions in A1~A4 are supported on this alternative pair of platforms and, thus, our findings regarding the first-review effect are robust to platform choice.

3.5.6 Temporal Price Variation

In our main study, we use static price data as one of the control variables. While the prior literature has taken the same approach (e.g., Li and Hitt 2008, Godes and Silva 2012), it is based on an implicit assumption that prices do not change over time or, at least, temporal price variations are not correlated with both independent and dependent variables at the same time. Although we do not believe that this assumption distorts our results,[17] it is still useful to relax the assumption and confirm the first-review effect while using dynamic prices as a control. To conduct this analysis, we collect the historical price data for both Amazon US and Amazon Canada from the website keepa.com, and merge them with the review data collected for the analysis in Section 3.5.5. However, keepa.com has only been in service since 2014 and, thus, if a product's first review was posted prior to that date, price data are not available. Hence, in our analysis, we include only 56 products for which we could match the price history data with our review data from both Amazon

[17] The results of the no-control specifications (see Section 3.5.2) suggest that the estimates of our primary interests are robust to the inclusion or exclusion of price. Moreover, using daily price data from both Amazon and Best Buy, we find that firms do not change their prices in response to newly posted product reviews (see Appendix B for details of this analysis).

US and Amazon Canada. In this sample, 32.1 percent of the products received inconsistent first reviews across the two platforms.

Using this data set, we estimate Equations (7) and (9) with the same set of control variables as used in the previous section (other than prices), based on the one-year observation window. We report the estimation results in the second column of Table 12 (valence) and Table 13 (volume). As can be easily seen, the signs of the coefficients of our interests are consistent with our theoretical predictions and thus support all the predictions in A1~A4. Therefore, the first-review effect holds regardless of the assumption on temporal price variation.

3.5.7 Platform-Specific First-Review Effects

Recall that we base our identification of the first-review effect on comparison of reviews for the same products with inconsistent first reviews across the two platforms. To implement this identification strategy, we have considered common coefficients for $\Delta FNegative$ and its interaction with $FDuration$ on the two platforms. In doing so, we implicitly assume that the magnitude of the first-review effect is identical across the two platforms. However, it is possible that the first-review effect could be different across platforms. Hence, we examine the sensitivity of our results to the assumption of identical vs. platform-specific first-review effects.

For this purpose, we replace β_1 and β_3 in Equations (6) and (8) with platform-specific parameters β_1^j and β_3^j and estimate the following equations:

$$\begin{aligned}\Delta AR_{it} = {} & \beta_0 + \beta_1^A FNegative_i^A - \beta_1^B FNegative_i^B + \beta_2\, ln(FDuration_{it}) \\ & + \beta_3^A FNegative_i^A \times ln(FDuration_{it}) - \beta_3^B FNegative_i^B \times ln(FDuration_{it}) \\ & + \mathrm{X}_{\mathrm{it}}^{\mathrm{A}}\Gamma^{\mathrm{A}} - \mathrm{X}_{\mathrm{it}}^{\mathrm{B}}\Gamma^{\mathrm{B}} + \varepsilon_{it}\ ,\end{aligned} \tag{10}$$

$$\begin{aligned}\Delta \ln(NR_{it}) = {} & \beta_0 + \beta_1^A FNegative_i^A - \beta_1^B FNegative_i^B + \beta_2 FDuration_{it} \\ & + \beta_3^A FNegative_i^A \times ln(FDuration_{it}) - \beta_3^B FNegative_i^B \times ln(FDuration_{it}) \\ & + \mathrm{X}_{\mathrm{it}}^{\mathrm{A}}\Gamma^{\mathrm{A}} - \mathrm{X}_{\mathrm{it}}^{\mathrm{B}}\Gamma^{\mathrm{B}} + \varepsilon_{it}\ .\end{aligned} \tag{11}$$

In Equation (11), as before, we impose $\beta_1^A = \beta_1^B = 0$, since, by definition, the number of reviews in the first period cannot be different between a positive and a negative first review. This model uses a different identification strategy: Identification of β_1^j and β_3^j now derives from cross-product variation within a platform, rather than from cross-platform variation of given products. Note that the product-specific effects are still controlled, but do not appear in the above estimation equations since they are differenced out. We estimate this model using all data considered thus far. We report the estimation results for the vacuum cleaner data on Amazon vs. Best Buy in Column 4 of Table 8 and Column 5 of Table 9; for the toaster data on Amazon and Best Buy in Column 4 of Tables 10 and 11; for the vacuum cleaner data on Amazon US and Amazon Canada with static prices in Column 3 of Tables 12 and 13; and for the vacuum cleaner data on Amazon US and Amazon Canada with dynamic prices in Column 4 of Tables 12 and 13.

As shown in these tables, we can replicate all our results with platform-specific first-review parameters. First, in the valence studies, the signs of all the relevant first-review effect parameters for both platforms are consistent with our predictions. Moreover, all the estimates are significant ($p < 0.05$), except that of the interaction term for Amazon Canada with dynamic prices, which is insignificant mainly due to the limited sample size. Second, in the volume studies, the signs of the interaction terms are also consistent with our theoretical prediction. The estimates are significant at the 0.01 level, or are at least marginally significant, except those for Best Buy, which are not significant, despite correct signs, most likely because the variation in the number of reviews is relatively smaller on Best Buy than on Amazon. Therefore, we can confirm across all studies that the first review has a significant influence on both the valence and volume of online WOM, regardless of whether it is common or distinct across different platforms.

4 Conclusion

Two key metrics of WOM, valence and volume, have been independently found to influence the sales of a product (Babić et al. 2016, Floyd et al. 2014, You et al. 2015). The central question of this paper is whether these two metrics are interrelated and, if so, what implications this interdependence might have for the management of online consumer reviews. Our theoretical model of consumer purchase decisions and the review updating process establishes the interdependence of valence and volume of online consumer reviews, which leads us to derive the powerful impact of the first consumer review. Specific predictions of the first-review effect are empirically validated. We summarize the main findings of the paper as follows:

- The valence and volume of online consumer reviews influence each other in their evolution paths, as a consequence of the conditional availability of online reviews.
- The valence of the first review positively affects the valence of the entire set of reviews. While this effect may decrease over the long term, it is persistently observed for an extended time period, even 36 months after the first review posting.

- The valence of the first review also positively affects the overall volume of reviews. Due to the positive feedback loop between sales and WOM volume, this effect grows stronger over time, engendering strong path dependence for online consumer reviews.

The results of our study have important theoretical and practical implications. First, our findings bring to the fore the significance of a single review. Specifically, we demonstrate both theoretically and empirically that a product's first consumer review powerfully influences the entire evolution path of online consumer reviews for that product and, thus, potentially determines its fate. Specifically, when the first review is negative, the commanding first-review effect could curb the chances of the product's market success, thus significantly hurting the seller. Our research findings increase awareness of this important first-review effect while also highlighting the need for scholars and practitioners to develop effective preemptive strategies to reduce the chances of a product receiving a negative first review, as well as management strategies that can soften the impact of a negative first review after the fact.

Second, our work brings attention to a potential weakness of online consumer reviews. Our results suggest that a negative first review dramatically reduces the overall number of reviews. Thus, with a negative first review, the body of consumer reviews for that product may become less informative, either because any biased information in consumer reviews remains uncorrected or because consumers may even lose the opportunity to learn about the product at all. This finding implies that consumer reviews could become less reliable as a source of information and as a communication channel, depending on the valence of the first review. In today's market, businesses are increasingly shifting their attention and investment from firm- to user-generated information. Our research points to the need to integrate these two types of information into an effective overall communication strategy.

Third, our results on the asymmetric influence of the positive and negative first reviews also have interesting implications for promotional review policies. Recent empirical studies find evidence that firms have a high incentive to post positive reviews of their own products but negative reviews for those of their competitors (Luca and Zervas 2016, Mayzlin et al. 2014). Our findings suggest that promotional reviews, if posted as the first review, are very influential. Although such reviews may be created unethically, they actually help to facilitate informative online WOM if they promote the firms' own products, since a positive first review leads to high

WOM volume. In this case, any upward bias in the initial promotional review would be corrected as more genuine consumer reviews are posted. However, posting such reviews first with the objective of harming a competitor can severely damage online WOM, since an unfavorable first review reduces sales and, thus, the number of subsequent reviews. Consequently, the initial downward bias derived from a negative promotional review is less likely to be corrected, and the online WOM may fail to offer sufficient product information, which could severely harm the viability of the product being reviewed. Hence, it is important to consider such asymmetry when an online retail platform develops its review management policy, especially its policy regarding promotional reviews.

Our paper makes unique contributions to the literature on review dynamics. Thus far, the literature has focused on the individual reviewer's motivation and its impact on the evolution of the review valence (e.g., Wu and Huberman 2008, Li and Hitt 2008, Godes and Silva 2012, Moe and Schweidel 2012). By contrast, we examine the dynamics of both valence and volume and, more importantly, consider their evolution jointly. This approach allows us to uncover the path dependence of both valence and volume from the very first review, which cannot be explained by considering the evolution of the valence independently from that of the volume. Our most novel finding, the increasing impact of the first review on WOM volume, can be explained by the conditional availability of the reviews as modeled in our paper, but not by individual reviewer motivation, as suggested in the literature.

Finally, this paper is not without limitations, which also point to a few interesting avenues for future research. First, in our study, we focus on demonstrating the first-review effect by choosing a product category whereby online WOM is likely to sway consumer choices. However, the first-review effect may be weak or even absent for some product categories. By studying such categories, one may find boundary conditions for the first-review effect as well as factors that either strengthen or weaken it. We leave this for future research. Second, in our model, we abstract away from firms' strategic decisions to influence reviews, given our objective of establishing a first-review effect. Future research might investigate the impact of a firm's actions concerning its own profits, consumer surplus, and the review-generation process.

Appendix A. Proofs

Derivation of the expected utility of the consumer (EU_t)

Based on the pdf of the normal distribution, we have

$$EU_t = 1 - \frac{1}{c}\int_{-\infty}^{\infty} e^{-(r_{t-1}+\eta_t-\tau x-p)} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{\eta_t^2}{2\sigma_t^2}} \, d\eta_t$$

$$= 1 - \frac{1}{c} e^{-(r_{t-1}-\tau x-p)} \int_{-\infty}^{\infty} e^{-\eta_t} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{\eta_t^2}{2\sigma_t^2}} \, d\eta_t.$$

Using the Jacobian transformation ($\eta_t = -y$), we can rewrite the above expected utility as

$$EU_t = 1 - \frac{1}{c} e^{-(r_{t-1}-\tau x-p)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_t} e^{\frac{-\left(y-\sigma_t^2\right)^2+\sigma_t^4}{2\sigma_t^2}} \, dy \, .$$

Since $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_t} e^{\frac{-\left(y-\sigma_t^2\right)^2}{2\sigma_t^2}} \, dy = 1$ always holds, we have

$$EU_t = 1 - \frac{1}{c} e^{-(r_{t-1}-\tau\cdot x-p-\frac{\sigma_t^2}{2})} \, . \quad \square$$

Derivation of the sales (s_t)

A customer purchases the product if and only if $EU_t = 1 - \frac{1}{c} e^{-(r_{t-1}-\tau\cdot x-p-\frac{\sigma_t^2}{2})} \geq 0$, which, with the normalization c = e, is equivalent to $x \leq \frac{1}{\tau}\left(1 + r_{t-1} - p - \frac{\sigma_t^2}{2}\right)$. Since $x \sim U[0,1]$, in each period, among M consumers arriving in the market, $M \times \frac{1}{\tau}\left(1 + r_{t-1} - p - \frac{\sigma_t^2}{2}\right)$ consumers will buy the product. Therefore, the total sales in period t is given as $s_t = \frac{M}{\tau}\left(1 + r_{t-1} - p - \frac{\sigma_t^2}{2}\right)$. □

Proof of Theorem 1

Part (a): Given $n_{t+1} = n_t + \delta \cdot s_{t+1}$ and $s_{t+1} = \frac{M}{\tau}\left(1 + r_t - p - \frac{\alpha(1-\alpha)}{2n_t}\right)$, we have, by the chain rule,

$$\frac{\partial n_{t+1}}{\partial r_t} = \frac{\partial n_{t+1}}{\partial s_{t+1}} \frac{\partial s_{t+1}}{\partial r_t} = \frac{\delta M}{\tau} > 0.$$

Part (b): Let ν_{t+1} denote the volume of the newly posted reviews: $\nu_{t+1} = n_{t+1} - n_t = \delta \cdot s_{t+1}$. Noting that $h_t = r_t n_t$, we can rewrite r_{t+1} as

$$r_{t+1} = \frac{h_{t+1}}{n_{t+1}} = \frac{h_t + \alpha \nu_{t+1}}{n_t + \nu_{t+1}} = \frac{r_t n_t + \alpha \nu_{t+1}}{n_t + \nu_{t+1}} = r_t + \left(\frac{\nu_{t+1}}{n_t + \nu_{t+1}}\right)(\alpha - r_t).$$

Then, $\frac{\partial r_{t+1}}{\partial \nu_{t+1}} = \frac{n_t}{(n_t + \nu_{t+1})^2}\ (\alpha - r_t) > 0$ is equivalent to $r_t < \alpha$. Hence the results follow. □

To prove the two propositions of the paper, we first state and prove the following lemma.

Lemma A1. $s_i^+ - s_i^- > 0, \forall i = 1, \ldots, t$.

Proof. We prove the lemma by mathematical induction. First, when $i = 1$, $s_1^+ - s_1^- = 1 > 0$. Second, suppose $s_n^+ - s_n^- > 0$ holds. Then, given $s_t^k = \frac{M}{\tau}\left(\frac{2\cdot\left(h_0^k + \alpha\delta \sum_{i=1}^{t-1} s_i^k\right) - \alpha(1-\alpha)}{2\cdot\left(1 + \delta \sum_{i=1}^{t-1} s_i^k\right)} + 1 - p\right)$, $(s_{n+1}^+ - s_{n+1}^-) - (s_n^+ - s_n^-) > 0$ is equivalent to $2\delta s_n^- + \alpha\delta(3-\alpha)(s_n^+ - s_n^-) > 0$, which holds since $s_n^+ - s_n^- > 0$, and implies that $s_{n+1}^+ - s_{n+1}^- > 0$. Therefore, $s_i^+ - s_i^- > 0$ holds for $\forall i = 1, \ldots, t$. □

Proof of Proposition 1

Part (a): Since $n_t^+ = 1 + \delta \sum_{i=1}^t s_i^+$, $n_t^- = 1 + \delta \sum_{i=1}^t s_i^-$, $h_t^+ = 1 + \alpha \cdot \delta \sum_{i=1}^t s_i^+$, and $h_t^- = \alpha \cdot \delta \sum_{i=1}^t s_i^-$, it is easy to see that $r_t^+ - r_t^- = \frac{h_t^+}{n_t^+} - \frac{h_t^-}{n_t^-} = \frac{h_t^+ \cdot n_t^- - h_t^- \cdot n_t^+}{n_t^+ \cdot n_t^-} > 0$ is equivalent to $1 + \delta \sum_{i=1}^t s_i^- > 0$, which always holds since $s_i^- > 0$ $(\forall i = 1, \ldots, t)$. Hence $r_t^+ > r_t^-, \forall t$.

Part (b): Note that $n_t^+ = 1 + \delta \sum_{i=1}^t s_i^+$ and $n_t^- = 1 + \delta \sum_{i=1}^t s_i^-$. Then, since $s_i^+ - s_i^- > 0$ $(\forall i = 1, \ldots, t)$, by Lemma A1, we have $n_t^+ - n_t^- = \delta \cdot \{\sum_{i=1}^t s_i^+ - \sum_{i=1}^t s_i^-\} > 0$. □

Proof of Proposition 2

Part (a): We prove $\Delta r_t > \Delta r_{t+1}, \forall t$ by showing $r_t^+ > r_{t+1}^+$ and $r_t^- < r_{t+1}^-$ in order. First, since $n_t^+ = 1 + \delta \sum_{i=1}^t s_i^+$, $n_t^- = 1 + \delta \sum_{i=1}^t s_i^-$, $h_t^+ = 1 + \alpha \cdot \delta \sum_{i=1}^t s_i^+$, and $h_t^- = \alpha \cdot \delta \sum_{i=1}^t s_i^-$, we have

$$r_t^+ - r_{t+1}^+ = \frac{h_t^+}{n_t^+} - \frac{h_{t+1}^+}{n_{t+1}^+} = \frac{\left(1 + \alpha\delta \sum_{i=1}^t s_i^+\right)\cdot\left(1 + \delta \sum_{i=1}^{t+1} s_i^+\right) - \left(1 + \alpha\delta \sum_{i=1}^{t+1} s_i^+\right)\cdot\left(1 + \delta \sum_{i=1}^t s_i^+\right)}{\left(1 + \delta \sum_{i=1}^t s_i^+\right)\left(1 + \delta \sum_{i=1}^{t+1} s_i^+\right)}$$

$$= \frac{\delta s_{t+1}^+ (1-\alpha)}{\left(1 + \delta \sum_{i=1}^t s_i^+\right)\left(1 + \delta \sum_{i=1}^{t+1} s_i^+\right)} > 0, \text{ and}$$

$$r_t^- - r_{t+1}^- = \frac{h_t^-}{n_t^-} - \frac{h_{t+1}^-}{n_{t+1}^-} = \frac{\left(\alpha\delta \sum_{i=1}^t s_i^-\right)\cdot\left(1 + \delta \sum_{i=1}^{t+1} s_i^-\right) - \left(\alpha\delta \sum_{i=1}^{t+1} s_i^-\right)\cdot\left(1 + \delta \sum_{i=1}^t s_i^-\right)}{\left(1 + \delta \sum_{i=1}^t s_i^-\right)\left(1 + \delta \sum_{i=1}^{t+1} s_i^-\right)}$$

$$= -\frac{\alpha\delta s_{t+1}^-}{\left(1 + \delta \sum_{i=1}^t s_i^-\right)\left(1 + \delta \sum_{i=1}^{t+1} s_i^-\right)} < 0 \ .$$

Therefore, we have $\Delta r_t - \Delta r_{t+1} = (r_t^+ - r_t^-) - (r_{t+1}^+ - r_{t+1}^-) = (r_t^+ - r_{t+1}^+) - (r_t^- - r_{t+1}^-) > 0$.
Part (b): Given $n_{t+1} = n_t + \delta \cdot s_{t+1}$, we have $\Delta n_t - \Delta n_{t+1} = (n_t^+ - n_t^-) - (n_{t+1}^+ - n_{t+1}^-) = -\delta(s_{t+1}^+ - s_{t+1}^-) < 0$, by Lemma A1. □

Proof of Claims in Section 2.1 (Claim 1) and Section 2.4 (Claim 2)

Claim 1 (Impact of Review Informativeness on the First-Review Effect) *The first-review effect holds irrespective of the informativeness of the reviews.*

[Proof] Given that our main analysis assumes informative reviews after the first review, it suffices to show the case where the reviews are uninformative. By definition, the reviews are uninformative if each review is drawn from the Bernouilli distribution with $\alpha = 0.5$. Hence, this becomes a special case of our main model, where the true quality is given as $\alpha = 0.5$. Since our propositions as well as Theorem 1 hold for any value of $(0 < \alpha < 1)$, the first-review effect continues to hold even when the subsequent reviews are uninformative. □

Claim 2 (Sales Implications of the First Review Informativeness). *The sales of a high-quality product (i.e., $\alpha > 0.5$) are expected to be higher following an informative first review than an uninformative one. However, the sales of a low-quality product (i.e., $\alpha < 0.5$) will be lower following an informative first review than an uninformative one.*

[Proof] By definition, the expected future sales of a product are given as $S_1 \equiv \alpha \sum_{i=1}^{n} s_i^+ + (1 - \alpha) \sum_{i=1}^{n} s_i^-$ if the first review is informative but as $S_2 \equiv 0.5(\sum_{i=1}^{n} s_i^+ + \sum_{i=1}^{n} s_i^-)$ if the first review is uninformative. Now, $S_1 - S_2 = (\alpha - 0.5)(\sum_{i=1}^{n} s_i^+ - \sum_{i=1}^{n} s_i^-) > 0$ if and only if $\alpha > 0.5$, since by Lemma 1, $s_i^+ > s_i^-$ holds for all i. Hence, the results follow. □

Appendix B. Supplementary Empirical Analyses

Possibility of Price Endogeneity

Throughout the paper, we have focused on the demand-side dynamics, while keeping the supply side fixed. However, it is possible that firms may change their prices in response to the valence and volume of consumer reviews. If that is the case, this dynamic pricing strategy might alter the magnitude of the first-review effect. However, whether or not in reality firms engage in dynamic pricing is an empirical question. Hence, in this section, we examine the relationship between product reviews and prices.

For this purpose, we collect daily prices in the vacuum cleaner category from both Amazon and Best Buy from July 9 to October 8, 2016. For completeness, we cover all of the vacuum cleaner models available on both Amazon and Best Buy on the first day of data collection, i.e., July 9 2016, which results in a total of 327 matched products. Using this data set, we run the regression of the price on the valence and volume of online reviews, together with other controls, given as follows:

$$\ln(\mathrm{P}_{\mathrm{ijt}}) = \beta_0 + \beta_1 \mathrm{AR}_{\mathrm{ijt}} + \beta_2 \ln(\mathrm{NR}_{\mathrm{ijt}} + 1) + \beta_3 \mathrm{NoReview}_{\mathrm{ijt}} + \gamma_{\mathrm{i}} + \epsilon_{\mathrm{ijt}},$$

where $\mathrm{P}_{\mathrm{ijt}}$ is the price of product i at time t in platform j, $\mathrm{AR}_{\mathrm{ijt}}$ is the average rating, $\mathrm{NR}_{\mathrm{ijt}}$ is the number of reviews, $\mathrm{NoReview}_{\mathrm{ijt}}$ is the dummy variable indicating no review was posted for the product, γ_{i} is the product-level fixed effect, and ϵ_{ijt} is the zero-mean error term following normal distribution. Note that, due to the product-level fixed effect, our estimation uses within-product variation in identifying the impact of the reviews on prices. We use OLS to estimate this model with clustered standard errors at the product level.

Table B1 reports the results. As easily seen from the table, in both Amazon and Best Buy data, none of the three coefficients of the product review variables is significant, indicating that there is no evidence that firms change their prices in response to online reviews. This result remain robust to an alternative specification with lagged terms (for AR, NR, and NoReview). These result justify our approach of focusing on demand-side dynamics while abstracting away from firms' pricing decisions.

Table B1. Parameter Estimates: Price Regressions

	Amazon		Best Buy	
	Estimate	S.E.	Estimate	S.E.
(Intercept)	4.164***	0.342	4.240***	0.029
AR	-0.029	0.045	0.001	0.007
$\ln(\text{NR} + 1)$	0.015	0.023	0.006	0.008
No Review Dummy	-0.279	0.369	0.033	0.033
Adjusted R^2	0.9709		0.9911	

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the product level. Product-level fixed effects are included.

Quality and the First Review

In our theoretical analysis, we examine the implication of the first review's informativeness by assuming both cases (informative and uninformative). However, whether the first reviews are indeed informative is an empirical question. For this purpose, we merge our data set for the main analysis with the quality ratings information collected from consumerreports.org as of March 6, 2015 (with an understanding that the consumer reports data reflect the perceived quality of experts rather than consumers). This results in a total of 128 products, which we use to calculate the correlation between product quality and the first-review rating. The correlation is given as 0.002 ($p = 0.9795$). Thus, it is hard to say that the first review correlates with product quality. However, the correlation between the average rating across all reviews and product quality is positive and significant ($\rho = 0.256, p < 0.01$), suggesting that the first review does not convey accurate product-quality information, probably due to a small sample size.

Review Text Analysis

In our theoretical analysis, we consider a product category in which quality carries more weight than fit in the purchase decision and reviews affect purchase decisions mainly through quality information. In our empirical analysis, we choose vacuum cleaners as such a category. Thus, in this section, we provide support for this choice by examining the review texts.

For this purpose, we first extract the one hundred most frequently used keywords from all the review texts associated with the review ratings that we use in our main analysis. We then ask three independent coders to classify each keyword into one of the four categories: horizontal attribute, vertical attribute, both, or neither. Note that coders were informed that these words are from the product reviews for vacuum cleaners.

First, none of the keywords is categorized as "both" by any coder, but thirty-five keywords are categorized as "neither" by all coders. All thirty-five are general keywords that have no specific relationship to the vacuum cleaner category (e.g., "bought," "I," "have," "just," and "one"). Hence, we continue our analysis with the remaining 65 keywords. Our analysis first shows that Cronbach's alpha, a measure of internal consistency, is 0.7, which gives us confidence as to their classification. Among the 195 evaluations (65 keywords × 3 coders), 74.4% fall in the "vertical attributes" category, while 2.5% are classified as "horizontal attributes." The proportion of each classification is significantly different from the others ($\chi^2(1) = 212.33, p < 0.01$). Table B2 provides examples of keywords. By our analysis, we support our choice of the vacuum cleaner category as one where quality information (rather than fit) plays an important role in consumer reviews.

Table B2: Examples of Most Frequently Used keywords in Consumer Reviews

Rank	Keywords	Frequency	The Proportion of coders who categorize the word as a vertical attribute
4	Easy	35502	1
5	Clean	32661	1
6	Well	29991	0.66
11	Carpet	26424	1
14	Suction	23572	1
16	Hair	22658	1
23	Floor	20157	1
28	Dirt	16870	1
36	Light	13722	1
37	Water	13556	0.66
38	Power	13403	1
44	Battery	12899	1
49	Cord	12250	0.66
51	Long	12157	0.66
57	Dust	11341	1
98	Tile	7243	1
99	Handle	7185	1
100	Charge	6917	1

Summary Statistics for the Data Sets used in Sections 3.6.4 and 3.6.5

Table B3. Summary Statistics (Toaster)

	Amazon		Best Buy	
	Mean	Standard Deviation	Mean	Standard Deviation
Proportion of negative first reviews	0.381	0.488	0.239	0.428
Average rating	3.634	0.649	4.074	0.885
Number of reviews	444.106	671.155	50.699	103.856
Price	70.790	69.125	75.512	71.274
Word count in the first review	138.027	150.173	46.257	48.140
Number of days since the first review	1645.027	1175.829	1013.115	709.829

Table B4. Summary Statistics (Vacuum Cleaners: Amazon US and Canada)

	Amazon US		Amazon Canada	
	Mean	Standard Deviation	Mean	Standard Deviation
Proportion of negative first reviews	0.287	0.454	0.184	0.389
Average rating	3.840	0.566	3.950	0.920
Number of reviews	946.287	1758.765	27.667	58.243
Price	163.182	143.823	266.590	539.362
Word count in the first review	235.081	297.912	84.098	155.797
Number of days since the first review	1520.431	885.416	736.569	541.503

Individual Review Level Analysis

In our analysis, we aggregate our individual review data to obtain monthly measures of WOM volume and valence. Since, by its nature, volume is an aggregate measure, aggregation is unavoidable. However, valence does not have to be aggregated and an individual-level analysis is possible. Thus, in this section, we examine whether we can observe the same first-review effect on valence, even at the individual review level.[18]

We use the same model as in equation (6) with the following modifications. First, the definition of FDuration is the number of days since the first review was posted. Second, we include product-level fixed effects to control product-level characteristics. With product-level fixed effects, we identify the first-review effect by exploiting the fact that some products start with different valences in their first reviews across the two platforms. Third, to account for differences

[18] We thank an anonymous reviewer and the Associate Editor for suggesting this analysis.

in rating behavior across the two platforms, we include a Best Buy platform dummy ($\mathrm{BB}_{\mathrm{it}}^{\mathrm{j}}$) and its interaction with $FDuration$. Finally, given the categorical nature of the dependent variable, we estimate the following ordered logit model and report the results in Table B5:

$$\mathrm{U}_{\mathrm{ir}}^{j*} = \beta_1 FNegative_i^j + \beta_2 FDuration_{ir}^j + \beta_3 FNegative_i^j \times FDuration_{ir}^j$$
$$+\beta_4 BB_{ir}^j + \beta_5 BB_{ir}^j \times FDuration_{ir}^j + \gamma_i + \varepsilon_{ir}^j$$
$$Rating_{ir}^j = 1 \Leftrightarrow U_{ir}^{j*} < \lambda_1$$
$$Rating_{ir}^j = k \Leftrightarrow U_{ir}^{j*} \in [\lambda_{k-1}, \lambda_k), where\ k = 2,3,4$$
$$Rating_{ir}^j = 5 \Leftrightarrow U_{ir}^{j*} > \lambda_4$$

where U_{ir}^{j*} is the consumer's latent utility of consumption for product i of review r in platform j; γ_i refers to the product-level fixed effects. The estimation results are consistent with our theoretical predictions (A1 and A3). First, the coefficient for $FNegative_i^j$ is negative and is statistically significant ($\beta_1 = -0.358, p < 0.01$). This suggests that, on average, the rating is lower when the first review is negative than when it is positive, thus supporting A1. Second, the coefficient of the interaction between $FNegative_i^j$ and $FDuration_{ir}^j$ is positive and is statistically significant ($p < 0.01$). This implies that the first-review effect on valence decreases over time, thus supporting A3. Finally, we reestimate the above equation with $FDuration$ replaced by Order (i.e., the order of the review in the sequence of reviews) as per Godes and Silva (2012). We obtain qualitatively similar results. Therefore, the first-review effect on the valence is robust to the disaggregation of product-review data.

Table B5. Individual Review Level Analysis: Valence of WOM

	Estimate	S.E.
FNegative	-3.58E-01***	1.18E-01
FDuration	-4.21E-04***	9.59E-05
FNegative × FDuration	6.97E-04***	1.46E-04
BB	2.28E-01***	6.25E-02
BB × FDuration	2.57E-04**	1.20E-04
Log likelihood	-59083.73	
AIC	118195.46	

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the product level. We combine the Amazon and Best Buy review samples for this analysis. Controlling for the review year fixed effects yields qualitatively the same results.

References

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of political Economy*, 113(1), 151-184.

Ameri, M., Honka, E., and Xie, Y. (2016). Word-of-Mouth, Observational Learning, and Product Adoption: Evidence from an Anime Platform. working paper.

Anderson, E. T., and Simester, D. I. (2014). Reviews Without a Purchase: Low Ratings, Loyal Customers, and Deception. *Journal of Marketing Research*, 51(3), 249-269.

Babić Rosario, A., Sotgiu, F., De Valck, K., and Bijmolt, T. H. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research*, 53(3), 297-318.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*,119(1), 249-275

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *The Journal of Economic Perspectives*, 151-170.

BrightLocal (2014), Local Consumer Review Survey, Survey Report, June 2014, https://www.brightlocal.com/wp-content/uploads/2014/07/Local-Consumer-Review-Survey-20141.pdf

Cameron, S. V., and Taber, C. (2004). Estimation of Educational Borrowing Constraints Using Returns to Schooling. *Journal of Political Economy*, 112(1), 132-182.

Chen, Y., Wang, Q., and Xie, J. (2011). Online Social Interactions: A Natural Experiment on Word of Mouth versus Observational Learning. *Journal of Marketing Research*, 48(2), 238-254.

Chen, Y., and Xie, J. (2008). Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. *Management Science*, 54(3), 477-491.

Chevalier, J. A., and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43(3), 345-354.

Chintagunta, P. K., Gopinath, S., and Venkataraman, S. (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), 944-957.

De Langhe, B., Fernbach, P. M., and Lichtenstein, D. R. (2016). Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, 42(6), 817-833.

Dimensional Research (2013). *Customer Service and Business Results: A Survey of Customer Service from Mid-Size Companies*, Research Report, April 2013, https://d16cvnquvjw7pr.cloudfront.net/resources/whitepapers/Zendesk_WP_Customer_Service_and_Business_Results.pdf eMarketer (2010), The Role of Customer Product Reviews, Research Report, January 2010, http://www.emarketer.com/Article/Role-of-Customer-Product-Reviews/1008019

Feng, J., and Xie, J. (2012). Research Note-Performance-Based Advertising: Advertising as Signals of Product Quality. *Information Systems Research*, 23(3-part-2), 1030-1041.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., and Freling, T. (2014). How Online Product Reviews Affect Retail Sales: A Meta-Analysis. *Journal of Retailing*, 90(2), 217-232.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets. *Information Systems Research*, 19(3), 291-313.

Gelman, A., and Hill, J. (2006). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Godes, D., and Mayzlin, D. (2004). Using Online Conversations to Study Word-of-Mouth Communication. *Marketing science*, 23(4), 545-560.

Godes, D., and Silva, J. C. (2012). Sequential and Temporal Dynamics of Online Opinion. *Marketing Science*, 31(3), 448-473.

Holmström, B., and Milgrom, P. (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, 55(2), 303-328.

Holmström, B., and Tirole, J. (1993). Market liquidity and performance monitoring. *Journal of Political Economy*, 678-709.

Ho-Dac, N. N., Carson, S. J., and Moore, W. L. (2013). The Effects of Positive and Negative Online Customer Reviews: Do Brand Strength and Category Maturity Matter? *Journal of Marketing*, 77(6), 37-53.

Kornish, L. J., and Li, Q. (2010). Optimal Referral Bonuses with Asymmetric Information: Firm-Offered and Interpersonal Incentives. *Marketing Science*, 29(1), 108-121.

Kuksov, D., and Xie, Y. (2010). Pricing, Frills, and Customer Ratings. *Marketing Science*, 29(5), 925-943.

Li, X., and Hitt, L. M. (2008). Self-selection and Information Role of Online Product Reviews. *Information Systems Research*, 19(4), 456-474.

Liang, K. Y., and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13-22.

Libai, B., Muller, E., and Peres, R. (2013). Decomposing the Value of Word-of-Mouth Seeding Programs: Acceleration versus Expansion. *Journal of Marketing Research*, 50(2), 161-176.

Liu, Y. (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing*, 70(3), 74-89.

Luca, M., and Zervas, G. (2016). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. *Management Science*, 62(12), 3412-3427.

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., and Pfann, G. (2013). More than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing*, 77(1), 87-103.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Review*, 104(8), 2421-2455.

Mizerski, R. W. (1982). An Attribution Explanation of the Disproportionate Influence of Unfavorable Information. *Journal of Consumer Research*, 9(3), 301-310.

Moe, W. W., and Schweidel, D. A. (2012). Online Product Opinions: Incidence, Evaluation, and Evolution. *Marketing Science*, 31(3), 372-386.

Muchnik, L., Aral, S., and Taylor, S. J. (2013). Social Influence Bias: A Randomized Experiment. *Science*, 341(6146), 647-651.

Ranard, Benjamin L., Rachel M. Werner, Tadas Antanavicius, H. Andrew Schwartz, Robert J. Smith, Zachary F. Meisel, David A. Asch, Lyle H. Ungar, and Raina M. Merchant. (2016). Yelp Reviews of Hospital Care Can Supplement and Inform Traditional Surveys of the Patient Experience of Care. *Health Affairs*, 35(4), 697-705.

Sun, M. (2012). How Does the Variance of Product Ratings Matter? Management Science, 58(4), 696-707.

Trusov, M., Bucklin, R. E., and Pauwels, K. (2009). Effects of Word-of-Mouth versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing*, 73(5), 90-102.

Wu, F., and Huberman, B. A. (2008). How Public Opinion Forms. C. Papadimitrou, S. Zhang, eds. *Internet and Network Economics*. Lecture Notes in Computer Science, Vol.5385. Springer, Berlin, 334-341.

You, Y., Vadakkepatt, G. G., and Joshi, A. M. (2015). A Meta-Analysis of Electronic Word-of-Mouth Elasticity. *Journal of Marketing*, 79(2), 19-39.

Zhang, J., and Liu, P. (2012). Rational Herding in Microloan Markets. *Management Science*, 58(5), 892-912.

Zhang, J., Liu, Y., and Chen, Y. (2015). Social Learning in Networks of Friends versus Strangers. *Marketing Science*, 34(4), 573-589.

Zhu, F., and Zhang, X. (2010). Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing*, 74(2), 133-148.

Tables

Table 1. Inconsistency in Online Reviews: Some Examples

Product		Average Rating			Number of Reviews			Price		
		Amazon		Walmart	Amazon		Walmart	Amazon		Walmart
A		4.0	>	2.2	303	>	4	79.00	=	79.00
B		2.5	<	4.0	8	<	81	36.68	<	39.80
C		4.2	>	3.0	328	>	2	80.99	<	89.99

Products (date when the data is recorded):
A=Canon Powershot ELPH 115 (6/6/2014); B=Toastmaster TM-103TR Stainless Steel 4 Slice Toaster Oven (4/16/2016); C=Sealy Baby Firm Rest Crib Mattress (4/13/2016)

Table 2. Theoretical Predictions

A1	The *First-Review* Effect on WOM valence	A product has a higher average rating after a positive first review than a negative first review.
A2	The *First-Review* Effect on WOM volume	A product has a higher number of reviews after a positive first review than a negative first review.
A3	The *First-Review* Effect on the dynamics of WOM valence	The advantage in the average rating due to the advantage from a positive first review decreases with time.
A4	The *First-Review* Effect on the dynamics of WOM volume	The advantage in the number of reviews due to the advantage from a positive first review increases with time (i.e., the Snowball Effect).

Table 3. Summary Statistics

	Amazon		Best Buy	
	Mean	Standard Deviation	Mean	Standard Deviation
Proportion of negative first reviews	0.254	0.437	0.181	0.386
Average rating	3.849	0.627	4.070	0.864
Number of reviews	380.655	673.972	58.226	104.865
Price	168.014	128.128	184.850	139.899
Word count in the first review	275.712	346.265	51.345	67.934
Number of days since the first review	1101.175	804.593	778.492	613.688

Table 4. Volume and Valence of WOM by the Valence of the First Review

First Review		Average Rating			Number of Reviews			
Amazon	Best Buy	Amazon	Best Buy	Difference	Amazon	Best Buy	Difference	Sample Size
(+)	(-)	3.974	2.835	1.138***	66.143	14.952	51.190**	21
(-)	(+)	3.046	4.264	-1.218***	41.147	18.970	22.176*	34
(+)	(+)	4.099	4.386	-0.286***	92.216	35.108	57.108***	111
(-)	(-)	3.151	2.857	0.293	55.182	17.545	37.636**	11
All Samples		3.823	4.083	-0.260***	77.011	28.525	48.486***	177

*p<.10, **p<.05, ***p<.01 (Paired t-test). This table is based on the paired item-level observations at the last month of the observation windows.

Table 5. Wilcoxon Signed-Rank Test

Period	Ranks for Average Rating			Ranks for Number of Reviews			Sample size
	First Review (+)	First Review (-)	Difference	First Review (+)	First Review (-)	Difference	
One Year	60.324	89.081	-28.757***	64.946	74.000	-9.054*	37
Six Months	75.870	113.739	-37.870***	77.174	91.652	-14.478**	46

*p<.10, **p<.05, ***p<.01

Table 6. Parameter Estimates: Valence of WOM

	Estimate	SrandardError
(Intercept)	1.135	0.777
ΔFNegative	-1.432***	0.259
$\ln(\text{FDuration})$	-0.130***	0.043
$\Delta\text{FNegative} \times \ln(\text{FDuration})$	0.498***	0.099
$\ln(\text{Price}^A)$	-0.228	0.288
$\ln(\text{Price}^B)$	-0.156	0.316
$\ln(\text{WC}^A)$	0.130**	0.066
$\ln(\text{WC}^B)$	0.113*	0.063
Type handheld	0.009	0.354
Type robotic	0.152	0.356
Type stick	-0.068	0.312
Type upright	-0.014	0.266
Bagged	-0.323	0.262
Corded	0.074	0.165
Search Product^A	-0.006	0.004
Search Product^B	0.014**	0.004
Search Brand^A	-0.007	0.006
Search Brand^B	-0.013**	0.006
Volume Proxy^A	0.253***	0.072
Volume Proxy^B	0.064	0.058
Adjusted R^2	0.2801	

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the product level.

Table 7. Parameter Estimates: Volume of WOM

	Estimate	Standard Error
(Intercept)	2.921***	1.045
FDuration	0.071***	0.011
ΔFNegative $\times$ FDuration	-0.047***	0.018
$\ln(\text{Price}^{A})$	-0.714**	0.290
$\ln(\text{Price}^{B})$	-0.266	0.277
$\ln(\text{WC}^{A})$	0.235***	0.076
$\ln(\text{WC}^{B})$	0.077	0.089
Type handheld	-1.207***	0.331
Type robotic	-0.187	0.461
Type stick	-0.448	0.303
Type upright	-0.669***	0.242
Bagged	-0.416*	0.217
Corded	-0.722***	0.263
Search Product^{A}	-0.004	0.008
Search Product^{B}	-0.010	0.007
Search Brand^{A}	0.017	0.019
Search Brand^{B}	0.035*	0.020
Valence Proxy^{A}	0.311***	0.102
Valence Proxy^{B}	0.032	0.110
Adjusted R^2	0.2873	

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the product level.

Table 8. Robustness Check: Valence of WOM

	(1) Observation Window: 2 years	(2) Observation Window: 3 years	(3) No Control Specification	(4) Platform-Specific First-Review Effect
(Intercept)	0.596 (0.845)	0.571 (0.804)	-0.052 (0.097)	0.751 (0.811)
ΔFNegative	-1.422*** (0.271)	-1.479*** (0.290)	-1.358*** (0.265)	
ln(FDuration)	-0.088** (0.041)	-0.062 (0.044)	-0.108*** (0.039)	-0.118*** (0.042)
ΔFNegative × ln(FDuration)	0.368*** (0.099)	0.392*** (0.103)	0.433*** (0.097)	
FNegative^A				-1.241*** (0.295)
FNegative^B				-1.681*** (0.331)
FNegative^A × ln(FDuration)				0.472*** (0.111)
FNegative^B × ln(FDuration)				0.524*** (0.146)
$\ln(\text{Price}^A)$	-0.137 (0.258)	-0.185 (0.274)		-0.178 (0.286)
$\ln(\text{Price}^B)$	-0.047 (0.262)	-0.097 (0.271)		-0.143 (0.319)
$\ln(\text{WC}^A)$	0.086 (0.064)	0.081 (0.062)		0.145** (0.066)
$\ln(\text{WC}^B)$	0.113* (0.066)	0.098 (0.063)		0.111* (0.061)
Type handheld	0.042 (0.306)	-0.061 (0.290)		0.091 (0.361)
Type robotic	0.442 (0.333)	0.389 (0.314)		0.098 (0.366)
Type stick	-0.062 (0.251)	-0.113 (0.240)		-0.111 (0.311)
Type upright	-0.040 (0.196)	-0.078 (0.187)		-0.022 (0.273)
Bagged	-0.471** (0.187)	-0.403** (0.172)		-0.322 (0.266)
Corded	0.298 (0.199)	0.301 (0.186)		0.081 (0.161)
Search Product^A	0.001 (0.004)	-0.002 (0.003)		-0.005 (0.004)
Search Product^B	0.010*** (0.003)	0.007*** (0.002)		0.014*** (0.004)
Search Brand^A	-0.006 (0.007)	-0.004 (0.006)		-0.010 (0.006)
Search Brand^B	-0.013** (0.005)	-0.010** (0.005)		-0.016** (0.006)
Volume Proxy^A	0.163** (0.067)	0.143** (0.063)		0.243*** (0.073)
Volume Proxy^B	0.036 (0.054)	0.034 (0.053)		0.057 (0.055)
Adjusted R^2	0.2753	0.2504	0.1342	0.2917

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the product level and are reported in the parentheses.

Table 9. Robustness Check: Volume of WOM

	(1) Observation Window: 2 years	(2) Observation Window: 3 years	(3) No Control Specification	(4) Display Effect	(5) Platform-Specific First-Review Effect
(Intercept)	3.517*** (0.981)	3.482*** (0.951)	0.661*** (0.115)	3.002*** (1.051)	3.291*** (1.014)
FDuration	0.051*** (0.007)	0.035*** (0.005)	0.073 (0.011)***	0.069*** (0.011)	0.085*** (0.013)
$\Delta\text{FNegative} \times \text{FDuration}$	-0.022* (0.011)	-0.014* (0.008)	-0.037* (0.022)	-0.049*** (0.018)	
$\text{FNegative}^{A} \times \text{FDuration}$					-0.077*** (0.023)
$\text{FNegative}^{B} \times \text{FDuration}$					-0.008 (0.027)
$\ln(\text{Price}^{A})$	-0.715** (0.313)	-0.651* (0.333)		-0.668** (0.284)	-0.796*** (0.291)
$\ln(\text{Price}^{B})$	-0.238 (0.290)	-0.192 (0.303)		-0.243 (0.271)	-0.306 (0.275)
$\ln(\text{WC}^{A})$	0.214*** (0.080)	0.203** (0.082)		0.237*** (0.079)	0.214*** (0.073)
$\ln(\text{WC}^{B})$	0.116 (0.087)	0.144* (0.083)		0.066 (0.087)	0.078 (0.087)
Type handheld	-1.110*** (0.370)	-1.049*** (0.384)		-1.212*** (0.328)	-1.248*** (0.327)
Type robotic	-0.100 (0.496)	-0.009 (0.512)		-0.222 (0.461)	-0.111 (0.442)
Type stick	-0.463 (0.326)	-0.300 (0.325)		-0.511* (0.301)	-0.353 (0.298)
Type upright	-0.704*** (0.251)	-0.614** (0.255)		-0.735*** (0.246)	-0.632*** (0.240)
Bagged	-0.290 (0.230)	-0.143 (0.230)		-0.374 (0.229)	-0.436** (0.216)
Corded	-0.664** (0.287)	-0.626** (0.280)		-0.728*** (0.248)	-0.719*** (0.259)
Search Product^{A}	-0.005 (0.006)	-0.002 (0.006)		-0.004 (0.007)	-0.006 (0.008)
Search Product^{B}	-0.007 (0.005)	-0.006 (0.005)		-0.008 (0.007)	-0.010 (0.008)
Search Brand^{A}	0.010 (0.016)	0.007 (0.016)		0.016 (0.017)	0.020 (0.020)
Search Brand^{B}	0.025 (0.017)	0.021 (0.016)		0.031* (0.017)	0.039* (0.020)
FirstPage^{A}				0.105 (0.217)	
FirstPage^{B}				0.486** (0.191)	
Valence Proxy^{A}	0.315*** (0.115)	0.314*** (0.117)		0.312*** (0.101)	0.322*** (0.101)
Valence Proxy^{B}	-0.002 (0.114)	-0.013 (0.109)		0.015 (0.110)	0.027 (0.110)
Adjusted R^2	0.2949	0.2898	0.0382	0.3115	0.2986

$^{*}p<.10$, $^{**}p<.05$, $^{***}p<.01$, Standard errors are clustered at the item level and are reported in the parentheses.

Table 10. Parameter Estimates: Valence of WOM (Toasters)

	(1) Observation Window: 1year	(2) Observation Window: 2years	(3) Observation Window: 3years	(4) Platform-Specific First-Review Effect
(Intercept)	-0.663 (0.995)	-1.300 (0.937)	-1.619^{*} (0.951)	-1.042 (1.031)
ΔFNegative	-2.120^{***} (0.295)	-2.048^{***} (0.284)	-2.056^{***} (0.289)	
ln(FDuration)	-0.118 (0.073)	-0.137^{**} (0.061)	-0.143^{**} (0.059)	-0.196^{**} (0.093)
ΔFNegative $\times$ ln(FDuration)	0.619^{***} (0.118)	0.416^{***} (0.106)	0.414^{***} (0.104)	
FNegative^{A}				-2.162^{***} (0.398)
FNegative^{B}				-2.036^{***} (0.477)
$\text{FNegative}^{A} \times \ln(\text{FDuration})$				0.745^{***} (0.171)
$\text{FNegative}^{B} \times \ln(\text{FDuration})$				0.431^{***} (0.155)
$\ln(\text{Price}^{A})$	0.862^{*} (0.466)	0.892^{**} (0.431)	0.935^{**} (0.449)	0.756 (0.469)
$\ln(\text{Price}^{B})$	0.642 (0.452)	0.636 (0.427)	0.630 (0.415)	0.478 (0.487)
$\ln(\text{WC}^{A})$	0.215^{**} (0.088)	0.249^{***} (0.084)	0.234^{***} (0.081)	0.206^{**} (0.091)
$\ln(\text{WC}^{B})$	0.131 (0.098)	0.094 (0.081)	0.079 (0.085)	0.102 (0.096)
Type toaster	0.306 (0.223)	0.249 (0.201)	0.124 (0.211)	0.361 (0.227)
$\text{Search Product}^{A}$	-0.002 (0.003)	-0.003 (0.003)	-0.002 (0.003)	-0.002 (0.003)
$\text{Search Product}^{B}$	0.012^{**} (0.006)	0.006 (0.004)	0.004 (0.004)	0.011^{**} (0.006)
Volume Proxy^{A}	0.147^{*} (0.081)	0.206^{***} (0.061)	0.195^{***} (0.058)	0.146^{*} (0.080)
Volume Proxy^{B}	0.278^{***} (0.091)	0.238^{***} (0.082)	0.248^{***} (0.075)	0.281^{***} (0.091)
Adjusted R^2	0.3980	0.4366	0.4158	0.4082

$^{*}p<.10$, $^{**}p<.05$, $^{***}p<.01$, Standard errors are clustered at the item level and are reported in the parentheses.

Table 11. Parameter Estimates: Volume of WOM (Toasters)

	(1) Observation Window: 1year	(2) Observation Window: 2years	(3) Observation Window: 3years	(4) Platform-Specific First-Review Effect
(Intercept)	0.942 (1.117)	0.829 (1.327)	0.208 (1.340)	0.991 (1.127)
$\ln(\text{FDuration})$	0.118*** (0.015)	0.050*** (0.010)	0.031*** (0.008)	0.131*** (0.022)
$\Delta\text{FNegative} \times \text{FDuration}$	-0.047* (0.027)	-0.038** (0.015)	-0.028*** (0.010)	
$\text{FNegative}^{A} \times \text{FDuration}$				-0.066* (0.035)
$\text{FNegative}^{B} \times \text{FDuration}$				-0.025 (0.037)
$\ln(\text{Price}^{A})$	-1.344* (0.718)	-1.433* (0.846)	-1.453 (0.891)	-1.297* (0.724)
$\ln(\text{Price}^{B})$	-1.190 (0.752)	-1.355 (0.908)	-1.528 (0.951)	-1.120 (0.768)
$\ln(\text{WC}^{A})$	-0.208 (0.154)	-0.115 (0.142)	-0.087 (0.137)	-0.196 (0.153)
$\ln(\text{WC}^{B})$	-0.210 (0.178)	-0.042 (0.192)	0.009 (0.193)	-0.215 (0.175)
Type toaster	0.324 (0.279)	0.334 (0.290)	0.239 (0.293)	0.309 (0.276)
$\text{Search Product}^{A}$	0.005 (0.006)	0.009 (0.006)	0.013** (0.005)	0.005 (0.006)
$\text{Search Product}^{B}$	0.007 (0.007)	0.004 (0.005)	0.003 (0.004)	0.007 (0.006)
Valence Proxy^{A}	0.472*** (0.141)	0.724*** (0.197)	0.871*** (0.178)	0.467*** (0.142)
Valence Proxy^{B}	0.385** (0.168)	0.435*** (0.156)	0.423*** (0.160)	0.382** (0.167)
Adjusted R^2	0.1899	0.2034	0.2142	0.1932

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the item level and are reported in the parentheses.

Table 12. Parameter Estimates: Valence of WOM (Amazon US and Amazon Canada)

	(1) Static Prices	(2) Dynamic Prices	(3) Platform-Specific First-Review Effect: Static Prices	(4) Platform-Specific First-Review Effect: Dynamic Prices
(Intercept)	-0.344	-0.803	-1.189	-1.788
	(0.992)	(1.830)	(0.754)	(1.282)
ΔFNegative	-1.736^{***}	-1.097^{***}		
	(0.244)	(0.338)		
ln(FDuration)	-0.094	-0.071	-0.021	-0.072
	(0.058)	(0.080)	(0.059)	(0.074)
ΔFNegative × ln(FDuration)	0.383^{***}	0.381^{***}		
	(0.103)	(0.138)		
FNegative^{U}			-0.947^{***}	-0.928^{***}
			(0.247)	(0.337)
FNegative^{C}			-3.015^{***}	-3.112^{***}
			(0.350)	(0.815)
FNegative^{U} × ln(FDuration)			0.283^{**}	0.461^{***}
			(0.119)	(0.161)
FNegative^{C} × ln(FDuration)			0.587^{***}	0.355
			(0.172)	(0.340)
$\ln(\text{Price}^{U})$	0.108	-0.048	0.263^{*}	0.191
	(0.199)	(0.291)	(0.143)	(0.235)
$\ln(\text{Price}^{C})$	0.030	-0.033	0.052	-0.050
	(0.160)	(0.267)	(0.119)	(0.267)
$\ln(\text{WC}^{U})$	-0.048	0.019	-0.035	0.003
	(0.065)	(0.072)	(0.052)	(0.065)
$\ln(\text{WC}^{C})$	-0.140^{***}	-0.003	-0.039	0.024
	(0.047)	(0.090)	(0.046)	(0.077)
Type hand	-0.197	-0.071	-0.009	0.151
	(0.385)	(0.398)	(0.005)	(0.369)
Type robot	0.160	-0.063	0.007	-0.318
	(0.378)	(0.566)	(0.004)	(0.540)
Type stick	-0.185	0.114	-0.001	0.539
	(0.388)	(0.547)	(0.008)	(0.505)
Type upright	-0.164	0.020	0.002	-0.126
	(0.311)	(0.392)	(0.006)	(0.387)
Bagged	-0.163	0.752^{*}	0.103	0.420
	(0.297)	(0.455)	(0.289)	(0.280)
Corded	0.170	-0.211	0.013	-0.074
	(0.202)	(0.328)	(0.325)	(0.286)
Search Product^{U}	-0.010^{*}	0.006	0.012^{*}	0.006
	(0.006)	(0.006)	(0.316)	(0.005)
Search Product^{C}	-0.009	-0.007	0.128	-0.005
	(0.006)	(0.006)	(0.266)	(0.005)
Search Brand^{U}	-0.001	-0.001	-0.040	-0.001
	(0.013)	(0.016)	(0.249)	(0.015)
Search Brand^{C}	-0.006	0.001	-0.177	0.006
	(0.011)	(0.012)	(0.156)	(0.009)
Volume Proxy^{U}	0.062	0.060	0.105^{*}	0.000
	(0.071)	(0.095)	(0.056)	(0.076)
Volume Proxy^{C}	0.020	-0.012	0.032	-0.002
	(0.084)	(0.119)	(0.081)	(0.107)
Adjusted R^2	0.3667	0.1992	0.5353	0.4158

$^{*}p<.10$, $^{**}p<.05$, $^{***}p<.01$, Standard errors are clustered at the item level and are reported in the parentheses.

Table 13. Parameter Estimates: Volume of WOM (Amazon US and Amazon Canada)

	(1) Static Prices	(2) Dynamic Prices	(3) Platform-Specific First-Review Effect: Static Prices	(4) Platform-Specific First-Review Effect: Dynamic Prices
(Intercept)	0.159	0.242	0.167	0.250
	(0.973)	(1.749)	(0.972)	(1.676)
FDuration	0.099***	0.077***	0.100***	0.068***
	(0.012)	(0.022)	(0.015)	(0.026)
ΔFNegative $\times$ FDuration	-0.053***	-0.080***		
	(0.019)	(0.024)		
$\text{FNegative}^{\text{U}}$ $\times$ FDuration			-0.056**	-0.063*
			(0.024)	(0.032)
$\text{FNegative}^{\text{C}}$ $\times$ FDuration			-0.049*	-0.141***
			(0.029)	(0.036)
$\ln(\text{Price}^{\text{U}})$	-0.147	-0.581*	-0.150	-0.584*
	(0.215)	(0.351)	(0.215)	(0.353)
$\ln(\text{Price}^{\text{C}})$	-0.200	-0.404	-0.201	-0.417
	(0.198)	(0.265)	(0.198)	(0.265)
$\ln(\text{WC}^{\text{U}})$	0.253***	0.110	0.252***	0.106
	(0.091)	(0.088)	(0.091)	(0.088)
$\ln(\text{WC}^{\text{C}})$	-0.048	-0.205**	-0.050	-0.194*
	(0.065)	(0.102)	(0.065)	(0.103)
Type hand	-0.055	0.472	-0.060	0.514
	(0.329)	(0.531)	(0.325)	(0.538)
Type robot	0.400	1.433	0.407	1.456*
	(0.466)	(0.882)	(0.472)	(0.868)
Type stick	-0.242	-0.615	-0.241	-0.518
	(0.384)	(0.517)	(0.383)	(0.544)
Type upright	-0.055	-0.304	-0.052	-0.309
	(0.251)	(0.337)	(0.252)	(0.331)
Bagged	-0.475	0.266	-0.479	0.261
	(0.311)	(0.369)	(0.312)	(0.366)
Corded	-0.004	0.625	-0.001	0.629
	(0.274)	(0.563)	(0.276)	(0.563)
Search Product^{U}	-0.014**	0.008	-0.014**	0.007
	(0.006)	(0.009)	(0.007)	(0.009)
Search Product^{C}	-0.000	0.009	-0.000	0.009
	(0.007)	(0.010)	(0.007)	(0.010)
Search Brand^{U}	0.031***	0.015	0.030***	0.015
	(0.012)	(0.018)	(0.012)	(0.017)
Search Brand^{C}	-0.008	-0.028**	-0.009	-0.029***
	(0.009)	(0.011)	(0.008)	(0.011)
Valence Proxy^{U}	0.292**	0.149	0.291**	0.152
	(0.137)	(0.201)	(0.137)	(0.201)
Valence Proxy^{C}	0.112	0.190	0.113	0.192
	(0.128)	(0.181)	(0.128)	(0.182)
Adjusted R^2	0.3466	0.3995	0.3463	0.4056

*p<.10, **p<.05, ***p<.01, Standard errors are clustered at the item level and are reported in the parentheses.