Measuring Product Type and Purchase Uncertainty with Online Product Ratings: A Theoretical Model and Empirical Application

Peiyu Chen,[a] Lorin M. Hitt,[b] Yili Hong,[c] Shinyi Wu[a]

[a] W. P. Carey School of Business, Arizona State University, Tempe, Arizona 85287; [b] The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [c] C. T. Bauer College of Business, University of Houston, Houston, Texas 77204
Contact: peiyu.chen@asu.edu (PC); lhitt@wharton.upenn.edu, https://orcid.org/0000-0003-1736-5711 (LMH); yilihong@uh.edu, https://orcid.org/0000-0002-0577-7877 (YH); shinyi.wu@asu.edu (SW)

Abstract. Building on the distinction between search and experience goods, as well as vertical and horizontal differentiation, we propose a set of theory-grounded, data-driven measures that allow us to measure not only product type (search vs. experience, horizontal vs. vertical differentiation) but also sources of uncertainty and to what extent consumer reviews help resolve uncertainty. The proposed measures have two advantages over prior methods: (1) unlike prior categorization schemes that classified goods as either search or experience goods, our measure is continuous, allowing us to rank-order the degree of search versus experience and horizontal versus vertical differentiation among products or categories. (2) Our approach is easier to implement than prior methods, because it relies solely on consumer ratings information (as opposed to expert judgment) and can be employed at multiple levels (attributes, products, or product categories). We illustrate empirical applications of our proposed measures using product rating data from Amazon.com. Our data-driven measures reveal the relative importance of fit in driving product utility and the importance of search for determining fit for each product category at Amazon. Our results also show that, while ratings based on verified purchasers are informative of objective product values, the current Amazon review system appears to have limited ability to resolve fit uncertainty. Our method and findings could facilitate further research on product review systems and enable quantitative measurement of product positioning to support marketing strategy for retailers and manufacturers, covering an expanded group of products.

1. Introduction

The internet has provided a convenient way for consumers to share their product and service consumption experiences through online platforms that curate consumer opinions. Over time, this collection of consumer evaluations has become a dominant source of information for consumers in determining their product choices. According to Dimensional Research, 90% of purchasing decisions are influenced by online reviews.[1] Scholars have utilized online product rating data to study the effects of online word of mouth on firms' strategies (Chen and Xie 2008, Kwark et al. 2014) and consumers' purchasing decisions (Dellarocas 2003, Li and Hitt 2010, Archak et al. 2011, Sun 2012). Whereas several studies have shown that product ratings affect sales, some studies have found no effects. The relationship between ratings and sales (Chevalier and Mayzlin 2006, Duan et al. 2008) and consumer satisfaction (Godes and Silva 2012) appears to vary across different product categories, indicating

that the efficacy of product reviews in informing consumers about product utility and resolving product uncertainty may vary by product characteristics.

Starting with Nelson (1970), prior research has typically classified products using a binary classification of search versus experience goods, and this has been the predominant approach in subsequent empirical work that needed to make this distinction. Based upon Nelson's (1970, 1974, 1981) seminal work, a search good is one for which consumers conduct a search to ascertain product utilities *before purchase* by means of *information search* or *direct product inspection* (e.g., trying on a dress). In contrast, when consumers rely on experience *after* purchasing the good (e.g., tasting canned tuna fish) to determine the utility of a good, it is classified as an experience good. Whereas this definition can be stated unambiguously, for the majority of the products, it remains a challenge to characterize a product as either a search or an experience good as a practical matter. It should also be

noted that the notion of search versus experience goods is related to, yet distinct from, the industrial organization concept of vertical versus horizontal differentiation, since experience goods (or search goods) can be vertically or horizontally differentiated. By definition, a vertically differentiated good is one for which the preference (among alternatives) is agreed upon (often termed the *common utility*). For example, a laptop with higher computing power is generally preferred over a laptop with lower computing power, all else being equal. A horizontally differentiated good is one for which the preference is not agreed upon (termed the *idiosyncratic utility*). For example, people prefer different colors of the same products. It is important to note that a product is often composed of multiple attributes, and these attributes may contribute to horizontal or vertical differentiation, and can be search or experience attributes—in forming overall utility, these attributes determine whether a product is horizontally or vertically differentiated and whether a search or experience good.

The internet has greatly facilitated searching of products and sharing of consumers' consumption experiences, potentially allowing consumers to evaluate product utility more accurately before purchase based on the experiences of other consumers. This has typically been described as reducing uncertainty in quality (the common utility), although some studies have suggested that experience can also help consumers assess idiosyncratic utility (due to fit or taste). To this general understanding, we note that increasing availability of information can shift the relative importance of vertical or horizontal attributes in purchasing decisions and allow consumers to more accurately evaluate product utilities before purchase, making experience goods behave more like search goods. For instance, if online consumers purchasing shoes value appearance (a horizontal, search attribute) as well as durability (a vertical, experience attribute), better information about durability gleaned from consumer reviews may shift choices to placing more weight on the vertical attribute (durability), even though the shoes themselves have not changed. Of course, these relationships likely vary by product, making the overall influence of reviews on the perception of product characteristics an empirical question.

In this research, we propose continuous, data-driven measures of product type, based upon the key definitions of search versus experience goods as well as vertical versus horizontal differentiation, that can be derived from archival product rating data. We also evaluate to what extent consumer reviews help resolve uncertainty before purchase. Our research is motivated by several observations: (1) that the binary classifications between search and experience goods and between horizontal and vertical differentiated

goods are often too coarse to reasonably describe complex differentiated products that dominate online and offline commerce in modern markets; (2) that online consumer reviews have facilitated discovery of product attributes and sharing of consumers' consumption experiences, potentially changing product types; and (3) that it would be substantially less difficult to construct product classification measures if they could be derived from widely available consumer review data, rather than product-by-product expert classification or indirect proxies (e.g., repair costs); and (4) that such measures may be more accurate because they are closely tied to consumers' information acquisition and decision-making processes. Finally, recognizing that the product-type classification is in many cases context-dependent and can be influenced by company strategy (e.g., design of review system or different ways to signal product value), our theory-grounded, data-driven method can be used at any level (category, brand, product, or attribute) and in any setting (e.g., online retail, auction, etc.) to measure product type within a particular context or market.

2. Related Literature
2.1. Online Product Reviews
Online product reviews contain information about products that could potentially affect purchasing behavior. A substantial number of studies link product reviews to market outcomes, such as product demand, for a variety of products sold online. Chevalier and Mayzlin (2006) found a significant effect of review valence on book sales, whereas Liu (2006) found that review volume is a better predictor of future sales than review valence. Other studies have suggested that variance of ratings affects the sales of some products, such as books (Sun 2012) and craft beer (Clemons et al. 2006). Further research has indicated that the identity of the reviewer provides additional information, moderating the effect of product reviews on sales (Forman et al. 2008). Taken collectively, these findings suggest that consumers find information in product reviews to be valuable in making product purchase decisions and that various statistics from online ratings and rating metadata provide useful information about product utility and consumer preferences. A related stream of literature examines the extent to which information in reviews may be biased and may not reflect true consumer utility. Dellarocas (2006) analyzed the potential of forum manipulation (firms pay reviewers to influence ratings) and concluded that, in equilibrium, this type of manipulation has a limited effect on review accuracy.

In our study, consumer reviews serve two roles: they contribute information to future customers, potentially increasing later consumers' ability to evaluate
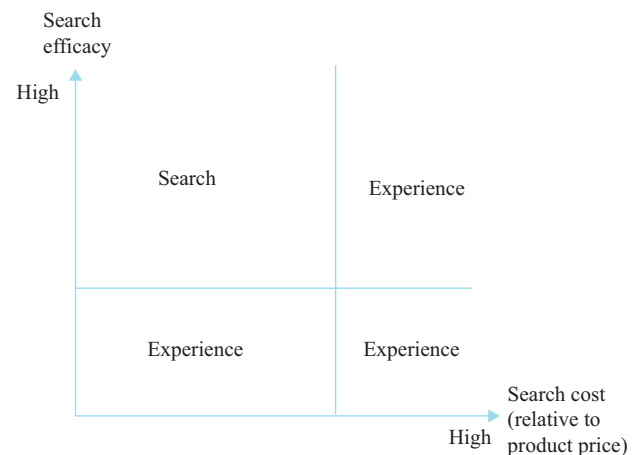
product utilities; and they provide information about consumer consumption experiences. We generally treat reviews as informative, although imperfect,[2] and rely on this information to make inferences about product characteristics. Our study is most closely related to analyses that connect the temporal evolution of online reviews to changes in consumer information. For instance, Godes and Silva (2012) attribute an observed downward trend in reviews for many products to the limited ability of consumers to assess large numbers of reviews causing increased uncertainty; Sun (2012) shows that the variance of reviews encodes useful information about products where reviews are mixed. In our analysis, we develop a theoretical model that ties certain statistics of reviews (mean and variance and their interactions) with product characteristics, which forms the basis of our measures to characterize products.

2.2. Product Type (Search vs. Experience Good)

The concept of "product type" originates from search theory (Stigler 1961), which is an economic framework for understanding how the search for price and product information influences product choices and pricing. The concept of product type was incorporated into the search theory by Nelson (1970, 1981) who distinguished two modes of consumer information acquisition: search and experience. By *search*, Nelson means "the direct inspection of a good, for example, a woman trying on a dress" (Nelson 1981, p. 43), and, by *experience*, he means "the determination of the utility of a good by means of experience after purchasing the good, for example, tasting canned tuna fish" (Nelson 1981, p. 43). Search efficiency determines whether consumers search to ascertain product utility prior to purchase or rely on experience through purchase to assess product utility and thus forms the distinction between *search* and *experience* goods. Specifically, when search cost is relatively low (in comparison with product price) *and* search is effective (i.e., it reduces uncertainty and helps consumers assess product utilities), consumers would find it worthwhile to "search" to ascertain product utilities before purchase, and these goods are classified as search goods (upper left quadrant in Figure 1). On the other hand, experience goods are those which consumers would rely on purchase to experience the product and determine utilities, because search cost (relative to product price) is too high or search efforts do not necessarily help consumers assess product utilities (e.g., checking a label of ingredients does not help assess "taste," and usefulness of software is difficult to be verified and ascertained with a description of functions).

Nelson operationalizes this distinction empirically by using product repair expenditures as a proportion

Figure 1. (Color online) Search Cost and Efficacy Determine Information Acquisition for Product Utility

of product sales for durable goods and determining whether sampling is destructive for nondurable goods (Table 2 and p. 320 of Nelson 1970). Nelson (1981) generalized his original model, allowing products to be described by both search and experience attributes, where product type is then determined by which type of attribute is dominant. The conceptual model in which products comprise both search and experience attributes implies a continuum of product types between the extremes of pure search and pure experience goods. However, empirical implementation of this construct remains challenging, and a dichotomous measure of product type is still in common use.

Ever since Nelson proposed the concept of product type, scholars have studied and linked this concept to different forms of market competition (Shapiro 1982; Villas-Boas 2004, 2006; Kwark et al. 2014), differences in marketing strategy (McWilliams and Siegel 2001, Hsieh et al. 2005, Park and Lee 2009), and consumer behavior (Klein 1998, Senecal and Nantel 2004, Weathers et al. 2007, Huang et al. 2009, Zhu and Zhang 2010, Mudambi and Schuff 2010, Sanders et al. 2016, Zhai et al. 2017, Li and Wu 2018). Overall, this literature clearly indicates the importance of understanding product type when assessing consumer behavior or designing a marketing strategy.

Despite greatly increased interest in search theory driven by the emergence of online commerce, the most widely adopted implementation of Nelson (1970)'s measure is still the coarse "repair expenditure" or "advertising expenses" metric from the 1960s or the classifications Nelson derived from these measures. This reliance has persisted, despite the fact that it is difficult to consistently and reliably classify products outside Nelson's original list (Mudambi and Schuff 2010), and the internet can dramatically change

consumers' information endowments, search costs, and search efficacy, altering the classification of search and experience products (Huang et al. 2009).

One response to this measurement difficulty is to use products that are seemingly located on extreme ends of the search-experience scale. For example, Senecal and Nantel (2004) contrasted calculators (an obvious search product) with wine (an obvious experience product) in studying the moderating effect of product type on the relationship between online product recommendations and consumer choice. Other alternatives are consumer surveys (Weathers et al. 2007), where consumers were asked about the importance of third-party information and direct personal experience in making an informed product choice. Other research used product prices as a proxy for product type, because higher-priced products have a greater return on search and are more likely to be search goods (Laband 1991). Nonetheless, although researchers have recommended a move from a dichotomous to a continuous search-experience classification (Nelson 1981, Klein 1998), empirical implementation remains a challenge, leaving scholars to continue relying on a dichotomous distinction for most empirical studies. Our approach provides an alternative data-driven and continuous measure of product type but maintains compatibility with existing approaches, since dichotomous measure is an extreme case of the continuous measure.

2.3. Product Differentiation (Vertical vs. Horizontal Differentiation)

There is a significant body of work on establishing a rigorous definition of product differentiation. Although scholars have offered different perspectives on the exact dimensions that define product differentiation, the existence of objective (common utility) and subjective (consumer-specific) components of quality has been two major sources of product differentiation. When the differentiation is based on the product attributes that most customers agree upon, it is defined as vertical differentiation. When the differentiation is based on attributes that depend on different consumer preferences or tastes, it is defined as horizontal differentiation. This distinction is most clearly articulated by Garvin (1984), who proposed a separation of quality assessment into an objective "product-based approach" and a subjective "user-based approach." In reality, a product may consist of both objective quality attributes with common utility to all customers and subjective attributes with idiosyncratic utility to different customers. For attributes that are commonly agreed upon, it is meaningful to aggregate the quality signals provided by consumers to obtain a single overall assessment of quality (e.g., a mean rating or

"number of stars"). However, for product attributes related to idiosyncratic consumer preferences, developing an objective quality measure that is independent of consumer preferences can be difficult and not meaningful (Theil 1971, Rosen 1974). Thus, instead of a priori separating objective and subjective qualities, we seek to take a consumer-specific utility-based view of product attributes where there are both horizontal and vertical attributes, with the overall product characteristics being determined by their relative importance.

3. Theoretical Model
3.1. Overview

As discussed before, product differentiation (horizontal vs. vertical) and product type (search vs. experience goods) are theoretical concepts that have long been established but difficult to conceptually disentangle and challenging to empirically measure because products often consist of numerous attributes[3] that may be horizontally or vertically differentiated, and can be best discovered by search or experience. For example, a pair of shoes may be evaluated by its material (vertically differentiated search attribute), look (horizontally differentiated search attribute), comfort (vertically differentiated experience attribute), or fit (horizontally differentiated experience attribute). Therefore, dichotomous measure of product differentiation and product type may be too restrictive. It is also important to note that the concepts of search and experience goods and vertical and horizontal differentiation are related but distinct. For example, chocolate milk and white milk are search goods and horizontally differentiated (some people prefer chocolate milk over white milk, while others do not), whereas certain cosmetic products are a better match for certain skin types than others, but this match is often difficult to ascertain without using the product—this would be a horizontal experience good (Li et al. 2011). In contrast, a digital camera with a high image pixel count is an example of a search attribute that is vertically differentiated, whereas hotel service quality is an experience attribute that is vertically differentiated.

The goal of this paper is twofold. First, we propose a data-driven method with continuous measures to help us understand where a product is positioned in the product type space (search vs. experience, and vertical vs. horizontal differentiation). Search versus experience reflects consumers' modes of information acquisition strategy to learn product utilities, whereas vertical versus horizontal differentiation reveals the relative importance of vertical versus horizontal attributes in affecting consumer utilities from the product. Second, we assess the degree of quality and fit

uncertainty faced by consumers before purchase and to what extent consumer reviews help resolve such uncertainty. Consumer reviews likely reduce consumer search costs and increase search efficacy, pushing more goods to fall into the upper left quadrant of "search goods" (Figure 1). However, information overload and presence of potential fraudulent reviews could also increase search costs and decrease search efficacy, making it more difficult to ascertain utilities before purchase. The effects are likely to be different across products. Understanding where a product is positioned in the product type space has important implications, as it allows firms to understand how their products are perceived by the consumers and to what extent consumers search to ascertain utilities before purchase, and thus helps inform or adjust marketing and pricing strategies. It is also important to know the efficacy of consumer reviews in resolving purchase uncertainty, because they have implications for company strategy, both in terms of how firms interact with review systems, as well as the way in which firms compete by engaging in product differentiation and informational advertising (Chen and Xie 2008).

Our measurement is motivated by a theoretical framework based on the research of the interactions among consumer reviews, consumer preferences, and product choice (Li and Hitt 2010, Godes and Silva 2012, Sun 2012, Kwark et al. 2014, Hu et al. 2017). Here, consumers make rational product choices based on available information (e.g., product descriptions and consumer reviews) and their search efforts to evaluate product features before purchase, and they then report their postconsumption evaluations in the form of product reviews and ratings. This data-generation process, wherein a consumer processes an *input* (reviews and other information via search) and produces an *output* (an additional review based on his/her consumption experience), provides information on consumer utilities and preferences; the time series of reviews can be used to understand the extent consumer reviews help resolve uncertainty and help later consumers assess products (Chen et al. 2018).

We consider a product as composed of multiple attributes, each of which may be horizontally or vertically differentiated. A common representation of the utility of multiattribute products is the "goods-attributes" approach, where the utility of the product is the summation of the utilities that a consumer obtains from each attribute (Lancaster 1966). Ex ante, the utility of each attribute may or may not be ascertained before purchase, the former being close to a search attribute and the latter being closer to an experience attribute. When some attributes cannot be assessed prior to purchase, consumers can only make purchase decisions based on the expected utility. Note that consumer reviews will *not* change whether a feature is

horizontally or vertically differentiated (as they are exogenous to consumer reviews); however, they may help (or hinder) consumers to evaluate product utility, which can change the relative importance of horizontal versus vertical attributes in consumers' choices. For instance, without reviews, it is difficult to make ex ante judgment of service quality (e.g., service quality during a hotel stay), but this information can be conveyed through consumer review causing a shift from experience to search, and given the fact that service quality is discoverable, consumers may place more weight on service, making the vertical attributes more salient. On the other hand, online sellers have the opportunity to "beautify" their product descriptions, exaggerating positive attributes while obscuring negative attributes, enhancing photographs, or fabricating information entirely, potentially making it more difficult for consumers to ascertain utilities before purchase. The design of the review system also affects the efficiency of information sharing among consumers. For example, some review systems allow consumers to provide contextual information (such as their body size) or rate products or services in different aspects, which may enhance the informational value of reviews in helping consumers evaluate product attributes. Thus, the role of reviews in shaping perceptions of product attributes is largely an empirical question and context-dependent. In the following section, we present a simple stylized model adapted from Sun (2012) to show how ratings and rating variance evolve under different scenarios and form the basis for our empirical inference and data-driven measures.

3.2. Model

Consumers' purchase decision and ratings of a product are influenced by the utilities from the product in question. Products often consist of multiple features or attributes, and each of these attributes may be horizontally or vertically differentiated. Following the standard goods-attributes approach (Lancaster 1966), we assume that utility is comprised of a weighted sum of the constituent attributes. For simplicity, we will use this structure to group all horizontal attributes into a single horizontal index of fit or taste (x), and all vertical attributes into a single objective quality index (q) in our consumer utility model. Consumers have the same appreciation of quality but different subjective utility due to different product taste or "fit." We assume that consumer fit (x) is uniformly distributed between [0, 1] and follows the linear Hotelling (1929) formulation with a misfit cost of t per unit of x. The parameter t is intrinsic to a product (category)[4] and dictates the importance of fit for consumers and therefore the degree of vertical versus horizontal differentiation, which we aim to uncover in our empirical application. Higher t implies greater horizontal differentiation and a

greater importance of fit. We also assume that t is known but product position relative to their ideal is not. This yields a consumer utility for a product of quality q with a price p for a consumer with a "distance" x from the product:

$$U(x,q) = q - t \times x - p.$$

Since price is a vertical attribute and is fully observable to consumers at the time of purchase, we can further combine q and p into a single measure, which we will refer to as *value* (i.e., the common utilities) to distinguish the contribution of vertical attributes from fit (as distinct from overall utility):

$$v = q - p.$$

Therefore, we can simplify the utility function to

$$U(x,v) = v - t \times x.$$

When consumers are uncertain about product value (or quality uncertainty, since price is known) and fit (fit uncertainty), both these constructs ex ante will be random variables ($\tilde{q}$ and $\tilde{x}$), which we treat as independent and based on information prior to purchase.

$$E[U(x,v)] = E[\tilde{v} - t \times \tilde{x}] = E[(\tilde{q} - p) - t \times \tilde{x}].$$

Consumers assess expected utility and decide to make a purchase if the expected utility meets or exceeds reservation utility (U_0). Consumers making a purchase will then rate the good,[5] which, for tractability, we restrict to being high (H), acceptable or medium (M), and low (L), which we will translate into a 1, 0.5, or 0 rating, respectively.[6] They will rate the product H if they are highly satisfied in the sense that the consumption utility exceeds reservation utility by a certain threshold (u), and rate the product M if the product is acceptable in the sense that the consumption utility meets reservation utility but does not exceed it by too much; they rate the product L if the consumption utility falls below reservation utility.[7]

Overall, our consumer decision process can be characterized as described in Figure 2, which is similar to Li et. al (2011) and Chen et. al. (2018).

3.2.1. No Uncertainty (Pure Search Goods). As a benchmark, we first derive the case when consumers are fully aware of product value and fit prior to purchase, which corresponds to a pure search good. Given a product offering a net value v and a consumer with fit x, the consumer will be able to make an optimal purchase decision under the following rule:

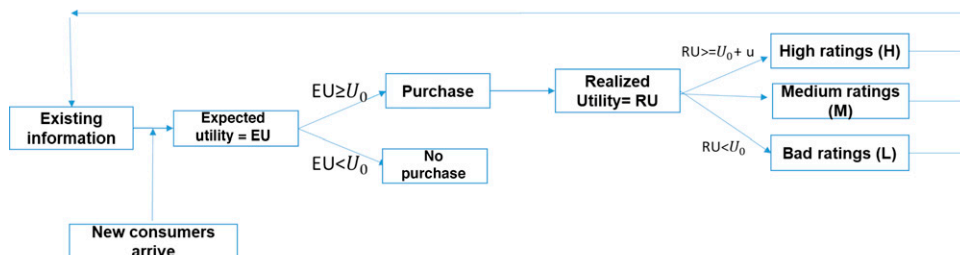If $E[U(x,v)] = U(x,v) = v - t \times x \geq U_0$, then purchase.

For notational simplicity, we will normalize reservation utility to zero ($U_0 = 0$). There are two corner solutions (*NU* denotes no uncertainty scenario). When $v < 0$, $D^{NU}(v) = 0$. This is because the product is "bad" in the sense that its product utility is worse than reservation utility for any x, and therefore no one would purchase it.

When $v \geq t$ (or, equivalently, $\frac{t}{v} \leq 1$), $D^{NU}(v) = 1$. In this case, t is small compared with v, and every consumer will purchase the product and be satisfied with the product regardless of fit, implying that there is no purchase uncertainty.

When $0 \leq v < t$ (or $\frac{t}{v} > 1$), fit will matter for purchase decision. We can find the consumer x_0 who is indifferent between buying and not buying: $v - t \times x_0 = 0$ or $x_0 = \frac{v}{t}$. All consumers to the left of x_0 will purchase, whereas all consumers to the right of x_0 will not purchase the product. Therefore, the demand of the product will be $D^{NU}(v) = x_0$, which increases in product value and decreases in t. Essentially, this is the region where fit plays a role for purchase decision. Further note that when $\frac{t}{v} > 2$ (or $< \frac{1}{2}t$), the product can be considered a niche product and will have small demand, as a majority of the consumers considering the product will end up not purchasing the product.

Given full information, the expected utility is actual consumption utility, and all consumers who purchase will give high or acceptable ratings to the product depending on their consumption utility (as those who would derive negative utilities from the product would not purchase). As we have noted, consumers will give high ratings if the product gives some

Figure 2. (Color online) Consumer Purchase and Rating Process

surplus utility at least by u on top of the reservation utility. Specifically, any consumer with $x \leq x_H$ will give the product high ratings (H), where

$$v - t \times x_H = u;$$
$$x_H = \frac{v - u}{t}.$$

Note that we assume u is not so high, so that at least some people are willing to give high ratings (i.e., $u < v$). Consumers $x_H < x \leq x_0$ will give the product acceptable ratings (M). The mean rating and rating variance of the product under no uncertainty will therefore be

$$r^{NU} = \frac{\left(\frac{(v-u)}{t} \times H + \frac{u}{t} \times M\right)}{D(v)} = \frac{\left(\frac{(v-u)}{t} + \frac{u}{2t}\right)}{\frac{v}{t}} = 1 - \frac{u}{2v}; \quad (1)$$

$$V^{NU} = \frac{\frac{v-u}{t} \times (1 - r^{NU})^2 + \frac{u}{t} \times (0.5 - r^{NU})^2}{\frac{v}{t}} = \frac{1}{4} \times \frac{u}{v} \times \left(1 - \frac{u}{v}\right).$$

Equations (1) and (2) show that when consumers can evaluate product utilities perfectly (they face no uncertainty), mean rating as well as rating variance are shaped primarily by $\frac{u}{v}$, the common (objective) utility (v), as well as the utility surplus required to give high ratings (u). The higher the common utilities offered by the product (e.g., when the product offers high quality and/or low price), the higher the mean rating will be, and the more demanding consumers are in giving high ratings (i.e., high u), the lower the mean rating will be. In addition, another important characteristic when consumers do not face uncertainty before purchase is that there will be no low ratings. We also note that the influence of market competition and consumer characteristics shall reflect on the common utility offered by the product as well as the surplus. When a product category has high market competition (which will likely drive up quality and lower price), the common utilities offered by the products in the category are likely to be high, and, therefore, ratings will be higher. In addition, when consumers are expecting higher surplus (u is high) in a given category, mean ratings would likely be lower than when consumers are more easily satisfied.

3.2.2. Introducing Uncertainty. Some goods are unambiguously describable such that product utilities can be correctly inferred, whereas others are not—a characteristic we term *feature observability*. Goods with high feature observability will behave as pure search goods. Goods where features cannot be easily observed or searched, and utilities cannot be ascertained before purchase will behave as experience goods, unless consumers have an alternative method to resolve this uncertainty. Consumer reviews have the potential to provide this information and resolve uncertainty, and the extent to which this is true we refer to as *review communicability*.[8] In this section, we examine

goods where there is quality or fit uncertainty (low feature observability) and the ability of reviews to resolve this uncertainty.

To capture the effect of information arriving over time due to reviews, we consider a two-period setting with a unit mass of consumers arriving in each period who purchase at most one unit of the good. In period 1, there are no reviews, so period 1 consumers make decisions based on the information they are able to gather. Those consumers that purchase the product can choose to provide a review, which is available to consumers in period 2 as additional information for decision making, and these consumers in period 2 can also provide reviews if they purchase. This two-period model enables us to study how consumers provide ratings in period 1 and examine the effect of period 1 ratings on later consumers' purchase decisions and subsequent ratings in period 2.[9] Whereas our model has only two periods, this process can theoretically continue indefinitely.

To model quality uncertainty, we assume that ex ante product value (as defined before as the difference between quality and price) is a random draw from a distribution with mean μ and variance σ_v^2, and that consumers are aware of the distribution and have rational expectations, even if they cannot discern the quality of a particular product before purchase. In essence, μ is the utility that a consumer can rationally expect based on all search attributes, whereas σ_v is the part of the "experience" utility related to quality that a consumer is uncertain about before purchase. This appears reasonable, as consumers are likely to be generally aware of whether product qualities or values are "all the same" or there is important variation among products within category.[10] The degree of quality uncertainty can therefore be measured by coefficient of variation ($\frac{\sigma_v}{\mu}$). The higher $\frac{\sigma_v}{\mu}$ is, the higher the degree of quality uncertainty. The actual quality and therefore value of the product (v) will be revealed after purchase. Fit uncertainty is modeled by treating x as a random variable as before. We will focus our results on the case where consumers are risk-neutral, as it simplifies the results, but these results generalize to the risk-averse case as well.[11] Therefore, the expected utility is $E[U(x,v)] = E[\tilde{v} - t \times \tilde{x}] = \mu - \frac{1}{2}t$, and consumers will make a purchase when $E[U(x,v)] \geq U_0$ (normalized to zero). To guarantee that a market exists, we require $\mu \geq \frac{1}{2}t$; otherwise, there are no purchases and no ratings. This condition shows that the expected utility (μ) must be high enough for consumers to even consider purchasing. The required mean utility increases with t.

After purchase and consumption, consumers learn the true value (and thus quality) of the product and how well the product fit them (i.e., v and x are

revealed), and the true consumption utility is realized. As before, consumers give high ratings (H) when the realized utility exceeds reservation utility by u (or $v - t \times x \geq u$), acceptable ratings (M) when $0 \leq v - t \times x < u$, and low ratings (L) when the realized utility falls below reservation utility ($v - t \times x < 0$). Depending on the parameters, there are three regions in the solution space (see Figure 3).

Two of these are corner solutions corresponding to situations where fit does not matter for purchase decisions. When the product is "bad" in the sense that the product value is worse than reservation utility ($v < 0$) (the far left in Figure 3), no consumer receives positive utility, and so the product has uniformly bad ratings with no rating variance. Similarly, if the product is "excellent" in the sense that $v \geq t$ (the right of Figure 3), all consumers will be satisfied regardless of x (since $v - t \times x \geq 0$ for any x), so no one will give a bad rating. Note also that when $t = 0$ (i.e., fit is not a concern at all) or when every consumer has the same x, everyone will receive the same utility from the product and give the same ratings, and so rating variance will be zero. This highlights the fact that rating variance is driven by different fit experiences across consumers ($t \times x$). Quality would not affect rating variance *within a product*, because all consumers experience the same quality of the same product. Essentially, *given a product of value v, within-product rating variance reflects the level of fit uncertainty faced by consumers when purchasing the product.*

The intermediate region $0 \leq v < t$ $\left(\text{or } 1 < \frac{t}{v}\right)$ has mixed reviews that include low ratings as people for whom the fit is good rate the product favorably (H or M), and those where the fit is poor do not (Figure 3). This is also the region where consumers are not sure ex ante if the purchase is a good one or not for them. If their x turns out to be smaller than x_0, then the product will be a good buy for them; if not, then it is a bad purchase (Figure 4). However, ex ante they do not know their x, and therefore cannot accurately determine if the purchase is a good buy, because they cannot assess product utilities before purchase. Essentially, *this is the region $0 \leq v < t$, where consumers cannot ascertain product utilities before purchase and which therefore falls into the classification of "experience goods."* We next derive the rating distribution for this region.

After purchase and consumption, the objective product value (v) and consumers' locations (x) are revealed. Given a consumer's realized x, the consumer experiences utility for the product of value v:

$$U(x, v) = v - t \times x.$$

Any consumer with $x \leq x_H$ will give the product high ratings (H), where $x_H = \frac{v-u}{t}$; for consumers $x_H < x \leq x_0$, they will give the product acceptable ratings (M), and consumers with $x > x_0 = \frac{v}{t}$ will give low ratings (Figure 4). The average rating of the product with value v will therefore be (H, M, and L are mapped to 1, 0.5, and 0):

$$r_1(v) = \frac{v-u}{t} \times H + \frac{u}{t} \times M + \left(1 - \frac{v}{t}\right) \times L = \frac{2v-u}{2t}$$

$$= \frac{v}{t} \times \left(1 - \frac{u}{2v}\right) = \frac{v}{t} \times r^{NU} < r^{NU}. \tag{3}$$

The variance is given by

$$V_1(v) = \frac{4v - 3u}{4t} - (r_1(v))^2 = \frac{4v-3u}{4t} - \left(\frac{2v-u}{2t}\right)^2$$

$$= \left[\frac{v}{t} \times \left(1 - \frac{v}{t}\right) \times (r^{NU})^2 + \frac{v}{t} \times V^{NU}\right]; \tag{4}$$

$$V_1(v) - V^{NU} = \left[\frac{v}{t} \times \left(1 - \frac{u}{2*v}\right)^2 - \frac{1}{4} \times \frac{u}{v} \times \left(1 - \frac{u}{v}\right)\right] \times \left(1 - \frac{v}{t}\right)$$

$$= \left(\frac{v}{t} \times (r^{NU})^2 - V^{NU}\right) \times \left(1 - \frac{v}{t}\right) > 0.$$

Figure 3. (Color online) Quality Uncertainty and Rating Distributions
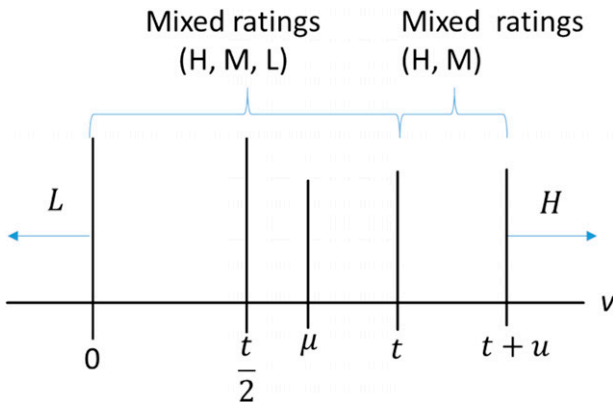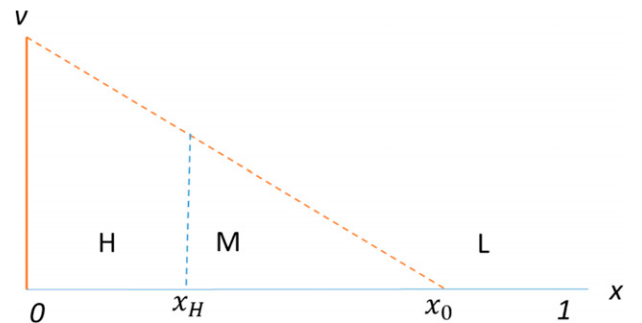

Figure 4. (Color online) Consumer Ratings When $0 \leq v < t$

Several observations are apparent from this model:

Observation 1. Equation (3) shows that ratings where consumers cannot ascertain product utilities before purchase (incomplete information) are lower than ratings without uncertainty (full information). Essentially, the difference in ratings reflects the effect of consumers' "inability" to determine product utilities before purchase and gives a direct definition of the value of "experience" in our context—*goods that have a significant experience component have a large difference between the full information outcome (as reflected in rating) and the incomplete information outcome.*[12] Equation (3) further indicates that such a difference can be measured as $\frac{t}{v}$. When t is high relative to objective product value (high $\frac{t}{v}$), consumers face higher uncertainty regarding whether the product will be a good buy for them, and they are more likely to rely on purchase to determine overall utility.

Observation 2. Figure 3 and Equation (3) show that mean ratings derived from purchasers are informative of product value or quality, when we compare goods that likely have the same t. Products that provide values lower than reservation utility ($v < 0$) will receive uniformly low ratings from their purchasers, and mean rating is increasing in product value or quality ($\frac{dr_1(v)}{dv} = \frac{r^{NU}}{t} > 0$).

Observation 3. Given the two Equations (3) and (4) and two unknowns (v, t), we can derive v and t:

$$v = \frac{2 \times r_1(v) \times t + u}{2} \; ; \tag{5}$$

$$t = \frac{u}{4 \times (r_1(v) - V_1(v) - r_1(v)^2)}. \tag{6}$$

This, in turn, provides information about the degree of horizontal differentiation (t) and the degree of experience versus search ($\frac{t}{v}$)—which reveals whether consumers could search to ascertain product utilities before purchase or simply rely on purchase to experience the product and reveal product utilities.

Observation 4. Quality uncertainty arises because different products offer different objective quality at a point in time in contrast with fit uncertainty, which arises because of differences in consumers' tastes (x) toward a product, which in turn generate variance in ratings within the same product. Note that quality uncertainty would affect product purchase decision but would not affect ratings variance within a product, because all consumers who purchase a particular product would experience the same quality for the same product as long as product attributes stay the same. Quality uncertainty, however, will be associated with variations in ratings among substitutable products (e.g., products in the same category), because, in the presence of quality uncertainty, some bad products may be purchased and receive bad ratings; if consumers are able to discern product qualities, then these bad products will not be purchased and rated. Therefore, the existence of rating variations across products in the same category is indicative of quality uncertainty faced by consumers.

To capture this effect, consider the coefficient of variation of quality ($\frac{\sigma_v}{\mu}$) within a set of products (a "category") as a candidate measure for quality uncertainty. This measure has desirable properties of a measure of quality uncertainty—when μ is very high, quality uncertainty will be low, because most products are "good buy" as they provide utilities higher than the reservation utility, so most products will receive high ratings from consumers (i.e., low ratings variation across products is expected); when σ_v is low, quality uncertainty will also be low, because products offer similar value μ, and therefore the average rating for each product (v_j) will also be similar, with little rating variation across products:

$$r_1(v_j) \approx \frac{\mu}{t} \times r^{NU}.$$

When $\frac{\sigma_v}{\mu}$ is high and consumers cannot discern product qualities, then consumers will shop randomly among products in a category and some consumers may end up purchasing a bad product that they should not have purchased. Thus, $\frac{\sigma_v}{\mu}$ captures the extent to which quality uncertainty can adversely affect consumers. We next show how we derive this measure:

$$V[r_1(\tilde{v})] = V\left[\frac{\tilde{v}}{t}\right] = \frac{\sigma_v^2}{t^2} \text{ (from (3))};$$

$$\sigma_v = SD[r_1(v)] \times t;$$

$$\frac{\sigma_v}{\mu} = \frac{SD[r_1(v)] \times t}{\mu}. \tag{7}$$

In sum, these results show that when there is no significant rating manipulation, purchasers' ratings provide useful signals regarding products' objective utilities and the degree of uncertainty faced by these consumers at the time of purchase. Whether these ratings are helpful or not for later consumers to ascertain product utilities depends on several factors that we will discuss next.

Period 2. So far, we have focused on consumer choice within a particular set of information conditions. To consider how changes in information influence customer choice, we must consider a multiperiod model. In period 2, a new group of customers arrive and demand at most one unit of the product. Consumers who arrive in period 2 can observe the mean ratings, rating variances, as well as textual reviews from earlier purchasers. These reviews may be helpful or not helpful in helping consumers ascertain product

utilities before purchase. *The second goal of this research is therefore to provide a theory-grounded empirical measure to assess to what extent reviews help potential purchasers ascertain utility before purchases.* We will discuss the rating outcomes that should happen when reviews are informative versus when they are not in helping consumers resolve uncertainty, and this will form the basis of our empirical application. Specifically, the degree to which ratings are trending toward the outcomes of reviews being informative measures the extent to which reviews help reduce purchase uncertainty for future consumers.

First consider the case that later consumers do not find or perceive earlier reviews to be informative in helping them ascertain utility from a product (i.e., uncertain remains); then the scenario considered in period 1 when consumers face uncertainty would still apply in period 2, and the results derived therein would continue to be relevant in the second period. There are several scenarios when consumers may not find reviews helpful. One is when product attributes change from one period to another (the most relevant case being the change of product price) that affects the objective utilities. Note that if consumers are aware of the exact changes (e.g., they know the adjustment of price), then consumers would still be able to correctly infer the common utilities that she could obtain from the purchase based on earlier reviews. In the case that consumers are not fully aware of the change (for example, consumers do not know whether the price has increased or not), they may not be able to accurately infer objective utilities from the product, and uncertainty about product utilities remains before purchase. Furthermore, when consumers perceive that there is rating manipulation or if self-selection bias exists in ratings (i.e., earlier reviewers tend to be more enthusiastic about the products and will provide more positive ratings), then again consumers may not perceive ratings to be helpful in ascertaining product utilities before purchases, and therefore the same results in period 1 where we introduce uncertainty would still apply (i.e., rating variance remains and ratings continue to be suboptimal).

We next consider the case when reviews from earlier purchasers are informative and help reduce purchase uncertainty and assist later potential purchasers to more accurately evaluate product utilities before purchase. We have shown that product ratings do carry signals of product quality/values when we compare ratings across similar products (Observation 2). Therefore, when product attributes do not change or when consumers are aware of the changes, consumers would be able to infer the common utilities of the product from its mean ratings, assuming there are sufficient number of reviews from period 1.[13] Consumers in period 2 can then update the expected value of the product according to existing ratings. In addition to learning objective common utilities

from ratings, consumers may also glean some information about product fit, such as (for example) examining reviewer metadata to identify reviews likely to have similar preferences (e.g., Li and Hitt 2008). We capture this effect by allowing consumers to learn their fit measure x with probability w.[14] In essence, w corresponds to the degree of review communicability described earlier. The more informative or communicative reviews are in conveying consumers' experiences that allow future consumers to assess their fit measure, the higher w is. Thus, w customers know their fit with certainty, whereas the remaining $(1 - w)$ remain uncertain about how well the product would fit them (it is expected in a multiperiod model that w would vary over time).

For the w consumers who can infer how well the product would fit them, they will make a purchase if and only if

$$E[U(x,v)] = E[\tilde{v} - t \times \tilde{x}] = v - t \times x \geq 0.$$

Again, we can find the consumer x_0 who is indifferent between buying and not buying:

$$v - t \times x_0 = 0;$$
$$x_0 = \frac{v}{t}.$$

So, consumers with $x \leq x_0$ will purchase the product, and these consumers will also give high or acceptable ratings to the product after consumption, accordingly to how well the product fits them.

The $(1 - w)$ consumers who remain uncertain about x will make a purchase if the expected utility exceeds the reservation utility:

$$E[U(x,v)] = E[\tilde{v} - t \times \tilde{x}] = E(v) - \frac{1}{2}t = v - \frac{1}{2}t \geq 0.$$

After consumption, consumers will rate the product based on whether their consumption utility exceeds, meets, or falls below their reservation utility (Figure 4). We derive the following results:

$$D_2(v) = w \times x_0 + (1 - w);$$

$$r_2(v) = \frac{\frac{v-u}{t} \times H + \frac{u}{t} \times M + (1-w)(1 - \frac{v}{t}) \times L}{w \times x_0 + (1 - w)}$$

$$= w \times \left(1 - \frac{u}{2v}\right) + (1 - w) \times \left(\frac{2v - u}{2t}\right);$$

$$\frac{\partial r_2(v)}{\partial w} = \left(1 - \frac{u}{2v}\right) - \left(\frac{2v - u}{2t}\right) = \left(1 - \frac{v}{t}\right) \times \left(1 - \frac{u}{2v}\right) > 0;$$

$$V_2(v) = w^2 \times V^{NU} + (1 - w)^2 \times V_1(v);$$

$$\frac{\partial V_2(v)}{\partial w} = 2 \times \left(w \times V^{NU} - (1 - w) \times V_1(v)\right) < 0. \quad (8)$$

These results suggest that rating variances would decrease and mean ratings would increase as review communicability improves (i.e., when reviews help consumers ascertain product utilities before purchase).

In sum, the aforementioned discussions suggest that, by comparing the rating dynamics across time, we can infer if reviews help consumers resolve uncertainty and ascertain product utilities before purchase. When rating variance remains and mean ratings do not improve, it would indicate that reviews do not help resolve purchase uncertainty. On the other hand, if we observe a decrease in rating variance and increase in mean ratings, then it would suggest that reviews are helpful in assisting consumers ascertain utilities before purchase.

Observation 5. As review communicability (w) increases, the rating variations from later consumers would decrease. This is because, when more consumers can learn their product utilities from existing reviews, there will be less "surprise" after consumption, and ratings are more likely to converge. Additionally, mean ratings will also increase when reviews become more informative for consumers to evaluate product utilities.

4. Empirical Application

Our theoretical model derived useful statistical properties related to product characteristics that form the basis of our identification strategy to characterize products. Specifically, given the same product, the common utilities are received by all consumers, and therefore the differences in utilities are driven primarily by consumers' uncertainty about fit arising from their idiosyncratic preferences, and this allows us to infer the degree of fit uncertainty and horizontal differentiation. Further, by comparing across products, differences in utilities include both the differences due to quality differences and consumers' idiosyncratic preferences, and when combined with the within-product variations, we can then infer a product's quality, with respect to others in the same category, and, based upon ratings across products, we can infer the quality uncertainty faced by consumers. Finally, by tracking rating dynamics over time, we can measure to what extent reviews help resolve uncertainty.

In this section, we apply our theory-grounded measures to empirically infer product type and estimate purchase uncertainty faced by consumers shopping at different categories at Amazon.com. The data required for our theory-grounded data-driven measures are readily available from most, if not all, rating systems. More specifically, we derive our measures by empirical observations of average product rating, temporal variance of product ratings, and within-category product variation. Collectively, these metrics will enable us to derive ordinal measures of product differentiation (t), degree of experience versus search (t/v), fit uncertainty, quality uncertainty (σ_v^2/μ), and review communicability (w). Whereas our focus is on the product category level for retail products, the same

approach is applicable at the product level or attribute level, provided reviews are available for the attributes, or other types of products or services (e.g., restaurants). Similarly, whereas we focus on simple category-level comparisons for demonstration purposes, these metrics can be potentially embedded in more complex models or estimated conditionally with additional controls for within-category variation or the underlying data-generation process of reviews.

4.1. Data

Using a combination of automated web crawlers and manual data gathering and cross-checking, we collected archival website data from Amazon for 15 predefined Amazon product categories from July 2004 to February 2012.[15] An observation in our data set is a particular review for a particular product and includes the rating and the time the rating was provided. To ensure a sufficiently large sample in a statistical sense (Stock and Watson 2003), we retain only products with at least 25 reviews. To maximize the number of products we consider while ensuring a relatively balanced sample, such that the key statistics (e.g., mean and standard deviation), we empirically compare across product categories are reliable, we perform the analysis on the first 55 reviews for each product. Table 1 provides the category-level summary of the data.

From these data, we can compute mean ratings at the product and category level, as well as the variance of these measures within products or categories or over time. Table 2 provides a category-level summary of our key statistics. To compute these ratings, we first normalize ratings by dividing by five, so that it is consistent with our theoretical model. We then compute a mean and standard deviation for each product (denoted by r_j and SD_j for the jth product within a category), which are then used to create category-level aggregates, as indicated in Table 2.

The mean rating (column 1) corresponds to ($r(v)$) in our theoretical model, and the average within-product rating variance (column 2) corresponds to ($V(r)$), which as discussed before, reflects the degree of fit uncertainty ($t \times x$) faced by consumers. Cross-products rating variations (column 3) measure $V[r_1(\tilde{v})]$ and will be useful in identifying quality uncertainty. The evolution of these metrics from period to period will enable the identification of w.

4.2. Measurement

Table 3 reports the estimates of the degree of horizontal differentiation (t) computed from Equation (6), the degree of quality uncertainty ($\frac{\sigma_v}{\mu}$) based on Equation (7), and the extent to which consumers rely on purchase to ascertain product utilities (i.e., experience good), $\frac{t}{v}$, based on Equations (5) and (6), for each

product category. We also visually represent the measurements in Figure 5 and 6. Figure 5 provides the position of each product category based on the experience versus search dimension (higher value means higher degree of "experience"—that is, consumers are less likely to conduct a search to ascertain product utilities before purchase and instead rely on purchase to learn product utilities) and product differentiation dimension (higher value indicates higher degree of horizontal differentiation), and Figure 6 shows the quality and fit uncertainty that consumers face when purchasing in each category.

To interpret the results, first consider the Fragrance category. This product is found to have the highest degree of horizontal differentiation (highest t) among all categories considered, which is consistent with the idea that preferences for fragrances are idiosyncratic, as consumers have different preferences for scents and the idiosyncratic utility is the major component of product utility. Whereas Fragrance is highly horizontally differentiated, our results show that most consumers ascertain the utilities of the product *before* purchasing fragrance products, rather than relying on purchase to determine utilities, at least in the Amazon context (Figure 5). This classification is consistent with Nelson's classification of clothing and footwear as search goods, because consumers are likely to conduct

search to ascertain utilities for these products before purchase (Nelson 1970, 1974). Examining the firms that participate in the Fragrance category at Amazon indicates that most products are "classical" products from well-known luxury brands and thus tend to have very high quality and therefore low quality uncertainty (as also revealed in Figure 6); the search cost to ascertain utility is also considerably lower for well-known brands. On the other hand, search efficacy is high when consumers engage in sampling for fragrance. Consumers could ascertain the utility of the product that they are purchasing by utilizing offline information or relying on prior experiences, for example, product samples that are often provided by well-known brands (e.g., through inserts in magazines), "test smell" of the fragrance in stores or from encounters with friends or others, or brands that they had purchased and experienced before. Relatively low search costs and high search efficacy put branded fragrances in the upper left quadrant "Search" in Figure 1. The market existence condition ($\mu \geq \frac{1}{2}t$) also explains why it makes sense that Amazon sells fragrance products from only well-known brands, because when expected utility (μ) is low, consumers would not consider purchasing at all. The measure is likely different if the market consists of less-known brands or products, where search costs to ascertain

Figure 5. (Color online) Degree of Experience vs. Product Differentiation Measures
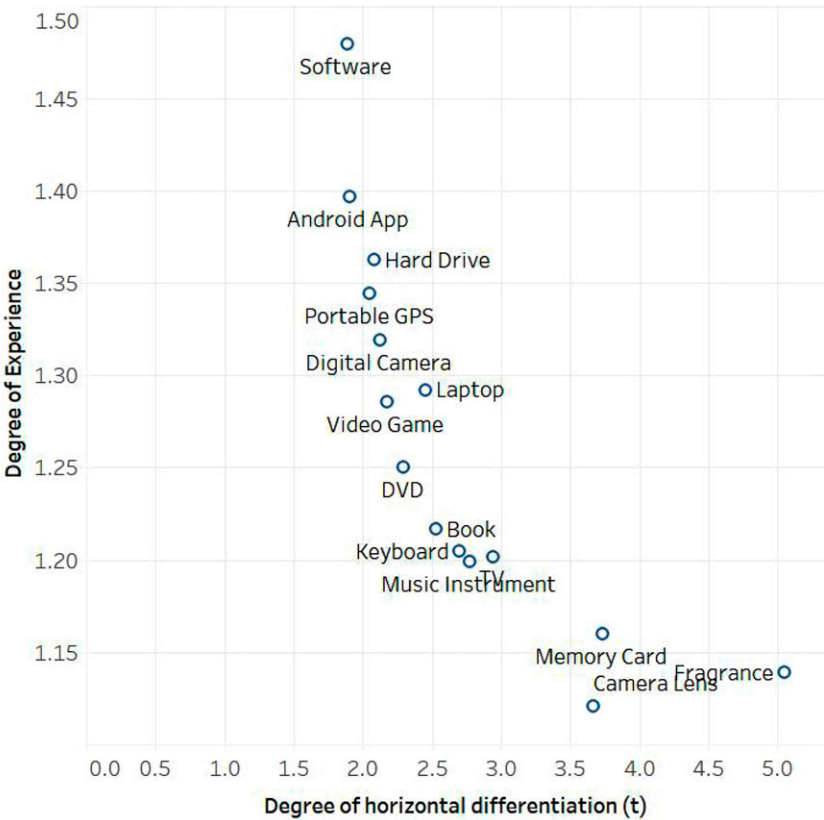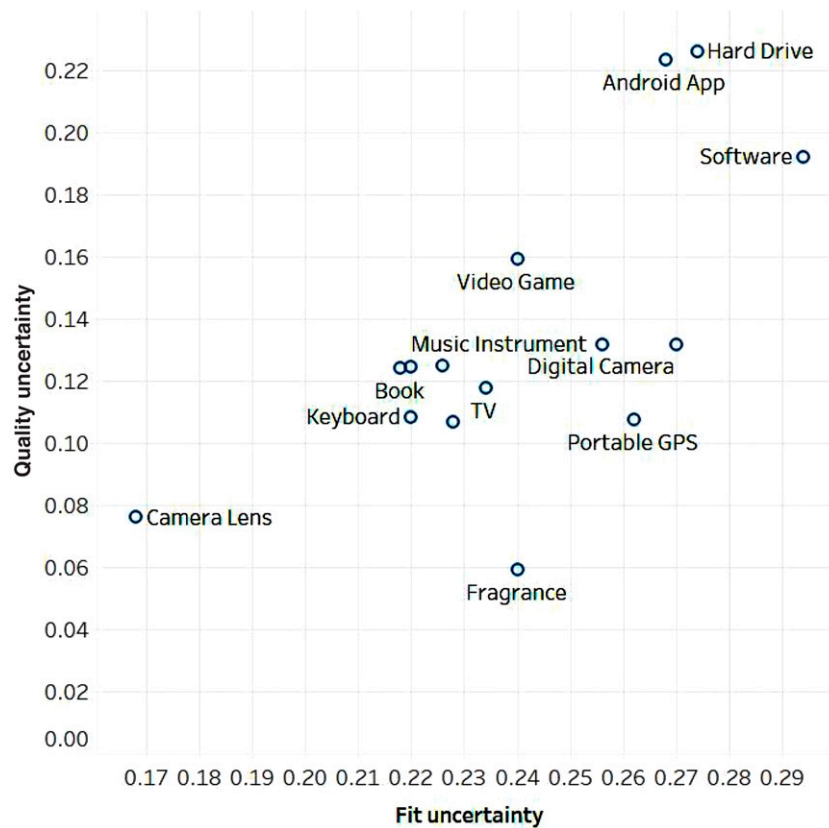
Figure 6. (Color online) Quality vs. Fit Uncertainty

utilities are presumably higher. Similar to the Fragrance category, Camera Lenses also have high t and low degree of "experience" (more "search good"), except the fit uncertainty is much lower than that for Fragrance, which is intuitive—it is much easier to describe and display a camera image online than it is to characterize an aroma.

We also note that, among all the categories considered, there is no category with very low t. Recall that the lower t is, the less important fit is in driving consumers' utilities (the less horizontally differentiated the product category is), and the more likely consumers agree on product utilities. These results suggest that products in general consist of nontrivial vertical and fit attributes. Among all categories, the Software and Apps categories have the lowest t. This makes sense, because individuals generally have similar requirements for software and apps, and therefore objective product utility remains important; however, some degree of horizontal differentiation is also present, since t is still substantial. We also see that, among all categories, Software and Apps have the most "experience goods," suggesting that it is most difficult to ascertain product utilities before purchase for these categories (Figure 5). The classification of Software and Apps as experience goods makes sense

and is consistent with the literature, because it is generally difficult to ascertain the utility of software without substantial use, and there is also wide range in attributes of these products that may be difficult to discern ex ante (e.g., "ease of use," "performance," etc.). That is, the search efficacy for software and apps is likely low, mapping to the lower left quadrant in Figure 1. Figure 5 shows that the Digital Camera

Table 1. Summary Statistics for Amazon Data Samples

#	Category	Number of products	Total rating observations
1	Android Apps	880	48,400
2	Video Game	648	28,751
3	TV	552	24,918
4	Digital Camera	473	20,285
5	Music Instrument	395	16,575
6	Books	384	15,889
7	Laptop	275	8,089
8	DVD	176	9,153
9	Camera Lens	152	7,146
10	Memory Card	151	6,726
11	Keyboard	137	6,900
12	Software	133	6,367
13	Portable GPS	98	4,945
14	Hard Drive	85	3,891
15	Fragrance	70	2,742

Table 2. Category-Level Rating Statistics

Category	Category-level rating mean: $avg(r_1, \ldots, r_J)$ (1)	Category-level average variation: $avg(SD_1, \ldots, SD_J)$ (2)	Cross-product rating variation: $SD(r_1, \ldots, r_J)$ (3)
Android App	0.716	0.268	0.160
Book	0.822	0.218	0.102
Camera Lens	0.892	0.168	0.068
Digital Camera	0.758	0.256	0.100
DVD	0.800	0.226	0.100
Fragrance	0.878	0.240	0.052
Hard Drive	0.734	0.274	0.166
Keyboard	0.830	0.220	0.090
Laptop	0.774	0.270	0.102
Memory Card	0.862	0.228	0.092
Music Instrument	0.834	0.220	0.104
Portable GPS	0.744	0.262	0.080
Software	0.676	0.294	0.130
TV	0.832	0.234	0.098
Video Game	0.778	0.240	0.124

category is somewhere in the middle in the measure of "experience good," suggesting that consumers face some level of uncertainty in determining product utilities. Interestingly, Nelson classified cameras as search goods in his 1970 paper (Table 2) but later recharacterized them as "experience goods" in his 1974 paper (p. 738), whereas Huang et. al. (2009, p. 67) suggested that classifying cameras as "search goods" could be more appropriate. Our finding that the Digital Camera category is somewhere in the middle may explain why such inconsistency arises and highlights the challenge of using a dichotomous measure.

A somewhat surprising finding is that consumers appear unable to, or do not attempt to, assess product utilities when purchasing in the Hard Drive category, at least in the Amazon context, and therefore it is classified as more experience-based. It would seem that hard drives can be characterized with a small set of dimensions (e.g., capacity, technology used, speed, etc.)

Table 3. Data-Driven Measurements

Category	Quality uncertainty	Degree of experience	t
Android App	0.223	1.397	1.901
Book	0.124	1.217	2.531
Camera Lens	0.076	1.121	3.670
Digital Camera	0.132	1.319	2.120
DVD	0.125	1.250	2.295
Fragrance	0.059	1.139	5.049
Hard Drive	0.226	1.362	2.080
Keyboard	0.108	1.205	2.697
Laptop	0.132	1.292	2.450
Memory Card	0.107	1.160	3.733
Music Instrument	0.125	1.199	2.776
Portable GPS	0.108	1.344	2.052
Software	0.192	1.479	1.886
TV	0.118	1.202	2.940
Video Game	0.159	1.285	2.172

and thus would be typically considered a search good—yet our results reveal that consumers face a high degree of both quality and fit uncertainty when purchasing hard drives on Amazon, making a hard drive purchase more of an experience good in practice, at least in the context of Amazon (Figure 6). These findings suggest that either existing information/reviews are not particularly informative of product utilities or consumers do not find it worthwhile to search/evaluate information to ascertain their utility before purchase. There are several reasons why a hard drive purchaser would rely on purchase more than search to ascertain utilities. First, the selection of hard drives is significantly more than that in traditional markets, resulting in much higher search/evaluation costs for buyers of hard drives in the online market. Second, for products such as hard drives, the objective measures of key attributes such as reliability may not be especially informative of actual experience (nearly all drives have a reported mean time between failure measured in decades, yet consumers routinely experience drive failures, and review texts also suggest many arrive "dead on arrival"). Consumption utility of a hard drive are also likely affected by the spin-up time, noise level, and performance of a particular drive, which are hard to ex ante evaluate and differ depending on how consumers use the drive. However, these issues likely do not surface until after consumption but will be reflected in consumers reviews of very different experiences, which contribute to uncertainty and low search efficacy. Given the huge selections of hard drive products online and the relatively high search costs when compared with the price of a hard drive (potentially positioning hard drives in the lower right quadrant in Figure 1), consumers may not expect much benefit from searching and therefore

likely forego search and rely on purchase to determine the utilities.

Overall, the previous discussions illustrate the usefulness of our data-driven approach, which allows us to measure key product characteristics (quality uncertainty, fit uncertainty, product differentiation, and search vs. experience) in the context where actual consumer choices are taking place. In many cases, these observations agree with prior ex ante classifications but can differ in potentially informative ways due to how the characteristics of goods and consumers interact in the online context. We also note that our data-driven approach can be applied at any level (category, brand, product, or attribute) and in any context, with ample opportunities for future researchers and practitioners to further improve upon precision of the metrics with more data, more observations of covariates, and more sophisticated empirical models.

4.3. Measuring Review Communicability

Over time, reviews reveal information, which may resolve quality uncertainty (as shown in Equation (3)). Further, according to Equation (8), when earlier purchasers' reviews help resolve fit uncertainty for later consumers, we expect lower rating variation from later purchasers, and the rate of decline in variance over time indicates to what extent fit uncertainty can be resolved by reviews (w) as reviews are accumulated.

To capture temporal variation in ratings, we follow Godes and Silva's (2012) approach of using the absolute deviation of the nth rating ($Deviation_{nj}$) from the prior average rating ($\overline{rating_{n-1,j}}$). This is more sensitive than the use of standard deviation (SD) for detecting changes and is less affected by general product-level standard deviation or any possible serial correlation in ratings:

$$Deviation_{nj} = \beta_j \times Sequence_{njc} + \alpha_{jc} + ym_{njc} + \varepsilon_{njc},$$

$$where\ Deviation_{nj} = abs\left(rating_{nj} - \overline{rating_{n-1,j}}\right). \quad (9)$$

In this equation, $Sequence_{njc}$ denotes the nth review for product j in category c, α_{jc} is the fixed effect for product j in category c, and ε_{njc} is the random noise associated with the nth review for product j in category c. In addition, we control for the year-month fixed effects (ym_{njc}) of each review to control for potential systematic differences of consumers arriving at different times. For example, holiday season consumers could be different from other consumers, or later consumers may have higher misfit cost (t) than earlier consumers, or vice versa. The aforementioned model was estimated using a panel-data fixed-effect approach. The results of the estimation are reported in Table 4. To interpret the estimation results, a larger number means the nth consumers' rating is more likely to deviate

from the prior average rating. From a statistical standpoint, this deviation also causes the sequence of standard deviation of the sample of first nth ratings to increase. In contrast, when reviews are effective in reducing uncertainty (w is higher), we expect that consumers will face less "surprise," which will cause ratings to converge (a negative β). Note that Equation (9) could be estimated at the product level or product-dimension level (if dimensional ratings are available) to understand review communicability at the product or product-dimension level. It is possible that experiences of some products or product features are easier to convey and communicate.

Based on results in Table 4, we can see that consumer reviews have significantly resolved fit uncertainty for purchases in the Memory Card, Laptop, Digital Camera, and DVD categories but not for other categories. The categories of Memory Card, Laptop, Digital Camera, and DVD appear to have a smaller set of well-defined product features, making it easier to communicate consumers' experiences of these features, compared with other categories that either have a wider set of product features or have certain important dimensions that are difficult to convey or be verified. It also appears that reviews do not help resolve fit uncertainty for categories that have high fit uncertainty and are mostly experience in nature (e.g., Software, Apps, and Hard Drive), because these categories all have features that seem difficult to convey or verify, for example, usefulness, ease of use, and complexity of installation for Software and Apps, and reliability for hard drives. Interestingly, we also find that fit uncertainty increases (though only marginally) for the Keyboard category as reviews accumulate, after controlling for the potential differences of reviews arriving at different times. This result suggests

Table 4. Panel Data Estimation Results of Amazon Data

Category	$\beta(Deviation)$	
Keyboard	0.0026*	
Music Instrument	0.0023	
Camera Lens	0.002	
Video Game	0.0004	
Android App	0.0002	
Hard Drive	0.0002	
Portable GPS	−0.0003	
Software	−0.0005	
TV	−0.0006	
Book	−0.0007	
DVD	−0.0017**	
Digital Camera	−0.0020***	
Fragrance	−0.0023	
Laptop	−0.0042***	
Memory Card	−0.0055***	Reviews resolve fit uncertainty

Notes. Estimation includes product-level fixed effects. Cluster robust standard errors are employed.
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

that additional reviews may have contributed to additional uncertainty for consumers purchasing in the Keyboard category. Overall, our findings suggest that, whereas the current rating system at Amazon can help resolve quality uncertainty, it remains limited in helping consumers resolve fit uncertainty for many product categories.

5. Discussion
5.1. Key Findings
In this paper, we develop measures that classify products along two types of distinctions: vertical versus horizontal differentiation, and search versus experience good, based on a stylized theoretical model of ratings. The key insight is that the variations in ratings within and across products and over time reveal important information about consumer preferences and consumer information at the time of purchase. We demonstrate the empirical application of the measures using a publicly available data set from Amazon.com. In addition to product type classification, we also show that we can identify the sources of uncertainty (quality or fit uncertainty) and to what extent consumer reviews help resolve uncertainty.

5.2. Strategic Applications
Generally, our metric has the advantages of being objectively defined, continuous, and applicable to any product that has a time series of consumer reviews. These advantages are especially useful in certain applications.

Determining Marketing Mix. The search versus experience goods measure has been previously used to identify appropriate marketing strategy (Klein 1998). Prior work has suggested that brand advertising is more effective for products with significant experience attributes, because the purpose is to boost expected utilities to induce purchases, whereas feature-based advertising can be more effective for search products. The form of product differentiation similarly informs which attributes should be emphasized in product promotions (Vandenbosch and Weinberg 1995, Makadok and Ross 2013). Our continuous measures may enable marketers to fine-tune these marketing tools, emphasizing attributes that are likely to differentiate products or resolve uncertainty. For instance, in products dominated by search attributes, there is a benefit of making investments along those dimensions (which may be one reason why technology products often exhibit "feature wars"). For products where the key attributes are difficult to perceive, sellers can focus on branding in the online context, perhaps coupled with alternative mechanisms (trial products, physical retail, advanced technology) to overcome the limited ability to communicate certain attributes in the traditional online context. One caveat here is that the search versus experience goods measure does not fully characterize a product, and, therefore, brand advertising may still be a viable marketing strategy for certain search products classified with our method.

Enhancing System Design. Our measures can also be used to enhance review system design. Our results have shown that the reviews based on verified purchasers are effective in resolving quality uncertainty, yet, the current Amazon review system has limited ability to reduce fit uncertainty. There appears to be an opportunity to improve review communicability, especially for products that contain horizontal experience attributes. Products for which choices are driven primarily by vertical attributes could be well covered by single-measure quantitative systems, such as single-dimensional star-rating systems (Chen et al. 2018). In contrast, systems that cover products with a high degree of horizontal experience attributes should consider ways of making additional information more readily accessible on both the attributes of the products and the attributes of the reviewers. Recent innovations that summarize review text are consistent with this approach.

In addition, as Villas-Boas (2004) has observed, consumers can gain further information regarding the fit of experience products to their preferences only by experiencing it after purchase. Therefore, for manufacturers and retailers, educating consumers about product attributes could lower prepurchase uncertainty and enhance performance. Online retailers can adopt state-of-the-art technologies such as a "3D virtual experience" (Jiang and Benbasat 2007) to simulate "touch and feel" experiences for potential consumers. A full understanding of what types of attributes play a dominant role in consumer decision is important for understanding the value of these technologies, because adopting such technologies is costly, both directly and in terms of opportunity costs of consumer attention. A successful implementation of the virtual reality technology should also make a product more search-based, and our measure provides a way to assess the efficacy of implementing the technology.

Engineering Social Influence. Products with different strengths on the search versus experience attributes spectrum may be subject to varying levels of social influence (Senecal and Nantel 2004, Huang et al. 2020). Because products dominated by search attributes benefit from the availability of information that helps consumers evaluate utilities they can derive from these attributes, consumers may be less likely to be persuaded by others who have had direct experience (social influence). However, for products dominated by

experience attributes, because consumers could not precisely estimate product utility without using it, social influence may become more important, as it might shape consumer preference or enhance expected utility to induce purchase. This will be an interesting future research question, and our proposed method provides several convenient and robust measures for researchers to include products of different nature as a construct.

5.3. Contribution

Our primary contribution is to build on the prior conception of product type (search vs. experience goods) and product differentiation (vertical vs. horizontal differentiation) and derive a set of theoretically grounded data-driven measures to classify products in greater precision (continuous vs. discrete) and also more objectively (using archival data rather than expert judgment). Most importantly, these measures will reflect (1) consumers' mode of information acquisition (search vs. experience) and whether reviews are an informative source for consumers to ascertain product utility before purchase, and (2) whether a product is dominated by horizontally differentiated versus vertically differentiated attributes. Whereas the definitions of these concepts would not change, the labeling of products can change due to different communication approaches, review system design, technology changes, or simple change of consumer behaviors and preferences, and, therefore, labeling of products is an empirical question that is highly context-dependent. Our data-driven measures provide a means to address these questions.

Given the extensive use of the Nelson (1970) framework for search versus experience goods in information systems, marketing, and management, the data-driven product-type measure potentially provides a much-needed update for the measurement of the product-type concept after almost half a century. In particular, it has been noted that the limitations of the original measurement scheme introduced uncertainty into the measurement of product type (Caves and Williamson 1985, Huang et al. 2009), especially for products that are outside of the original list (Mudambi and Schuff 2010). These problems are particularly acute for the study of online commerce, where there is a vast expansion in product variety and where the widely adopted "repair cost" measure is likely to be less relevant for many common products, such as digital content and other information goods. Further, our framework also allows us to measure "effective" product differentiation based on consumers' reviews that reflect the relative importance of horizontal attributes versus vertical attributes in driving consumption utilities. This adds objectivity of the vertical versus horizontal differentiation measure, which is traditionally done using market research, namely, surveys, based on a predefined set of characteristics. Consumer reviews likely reveal new features/attributes that companies may not take into account initially and provide a better and more accurate measure of product differentiation as perceived by consumers.

Our study also contributes to the ongoing stream in information systems that demonstrates the value of review data for answering fundamental questions of interest to information systems scholars. In addition to supporting the extensive literature on reviews and sales, our approach is similar in spirit to other works that have developed frameworks to extract otherwise unavailable information from online reviews, such as product characteristics that are difficult to describe (e.g., Ghose and Ipeirotis 2011) and the papers that followed that approach. We take advantage of the unprecedented availability of large-scale rating data that provide opportunities to measure and empirically classify products as perceived by consumers, thus closing the gap between theory and key measurements of product type and product differentiation. In addition to being a continuous measure (allowing us to compare across products), our data-driven method can adapt to changing product positions, information conditions, or consumer preference, since it does not rely on any sort of ex ante product classification. For example, as reviews become more informative, products can become more search-oriented, even for horizontally differentiated experience products, if consumers are able to find relevant information by learning from consumers with whom they share a similar preference.

As noted earlier, our approach supports the use of product classification for the development of marketing strategies or the improvement of recommender systems. For example, the method can be used for the decision of when to focus on brand rather than attribute marketing. The classification can also be used to determine when the review systems can efficiently use a simplified presentation (a single numeric scale), a more complicated quantitative scale (distinguishing vertical and horizontal attributes), or a richer presentation that enables the inference of reviewer preferences and the communication of complex attribute information (pictures, text, models, etc.). This is especially important given that consumers may have limited ability to process statistical information for complex choice tasks (Nisbett et al. 1983, Darke et al. 1998). Therefore, it is potentially helpful to limit information presented to consumers when it is unlikely to improve decision accuracy (Simon 1982, Jones et al. 2004). Online shopping is one scenario where consumers have access to a potentially large range of information, and only a small subset may be needed to attain high choice accuracy for at least some products. For

products where the full range of information is needed, it is still useful to understand what types of information can be aggregated (e.g., restaurant attributes that are primarily search) and what cannot. It may also be helpful in identifying when review systems are generally limited and when other types of approaches (discussion forums, product visualization tools) may be more effecting in aiding consumer choice.

5.4. Limitations and Opportunities for Future Research

Our study has a number of limitations, which present opportunities for future research.

First, this study uses the quantitative review component for measurement, and the assumption is that rating reflects consumers' overall satisfaction with their purchase. Consumers may also leave very detailed information in the review text, and those review texts are incorporated in our model as additional information sources for consumers' purchase decisions (as reflected in Figure 2). Review texts are also an important component in our model, as they serves as important mechanisms by which consumers can resolve uncertainty. Whereas our model allows us to infer review communicability (w), further research is needed to understand what factors contribute to review communicability and in what way consumers learn from others. For example, do consumers learn only from consumers of similar preferences, or does learning also take place across consumers of different preferences? A further investigation of textual reviews may help shed light on how consumers learn different product features. There is also potentially an opportunity to think about mapping textual reviews to quantitative information that would enable additional measurement to refine these models, and progress in that direction may also yield insights into how review systems can productively resolve product uncertainty by reducing the consumer burden of extracting information for textual or visual review material.

Second, the beauty of our proposed method is that it is theoretically grounded and has minimal data requirements, and the empirical applications we presented serve as a demonstration of the general framework that other academic scholars or industry analysts could use to infer product type or product differentiation, and, therefore, we did not include an exhaustive list of context-dependent control variables. Our hope is that this will prove useful as a basis for improving empirical work that requires the measures of constructs like vertical or horizontal differentiation or search versus experience components. Researchers can enhance measurements by adding additional controls (e.g., product or reviewer characteristics) or by improving context-specific choices, such as which reviews to include or exclude and what methods to use

to aggregate reviews over time and across products. We show that, even without adding additional control variables, our model provides reasonable estimates and produces plausible results. Future research could improve upon the empirical model by employing other identification strategies to achieve more precise estimation.

6. Conclusion

Data-driven analytics has attracted significant attention from both academics and practitioners. The digital revolution and digital innovation have generated data that researchers have only dreamed of a decade ago (Brynjolfsson and McAfee 2011). One challenge that remains is how to make sense of and leverage the freely available user-generated data to inform research and practice by building theoretically grounded and empirically validated analytical tools. This challenge has already caught the attention of many scholars in the field of information systems. As an important type of consumer-generated content, online product reviews provide a prime example of microlevel data that are free and readily available in digital format. A better understanding of the types of information that can be extracted from these data has considerable value. Our proposed theoretically grounded data-driven method shows that the dynamics of longitudinal ratings provides an effective way to understand consumer preference and their (in)ability to ascertain product utility before purchase, which in turn allows us to infer the relative importance of search versus experience attributes and vertical versus horizontal attributes in driving consumers' utilities. Understanding product type and product differentiation has important implications to marketing and sales strategy, product design, and website design, as well as to helping researchers measure these important constructs in their research studies.

Endnotes

[1] See http://www.reviewpro.com/survey-zendesk-mashable-dimensional-research-90-say-positive-reviews-impact-purchase-decisions-26016.

[2] It is well understood that reviews can include both bias and error due to deliberate manipulation (Dellarocas 2006, Luca and Zervas 2016), preference variations between reviewers and new purchasers

(Li and Hitt 2008), or dynamic effects as reviewers interact with each other (Moe and Schweidel 2012, Muchnik et al. 2013, Hu et al. 2017, Wang et al. 2018). However, consumers have mechanisms to address these effects, such as considering reviewer metadata (Forman et. al. 2008) and review text (Ghose and Ipeirotis 2011). It is also possible to limit the influence of biased reviews by disregarding reviews that are most likely to be biased (e.g., early reviews and reviews from nonpurchasers), and some review platforms also take measures to combat potential fraudulent reviews (e.g., verified purchases, label of incentivized reviews, and review filtering) and limit the visibility of reviews deemed to be of low quality (Luca and Zervas 2016).

[3] Here and throughout, we use the term *attribute* to describe product characteristics that affect the consumer. For our purposes, this is synonymous with the concept of a product *feature*, except that attribute does not necessarily imply a positive utility weight. For instance, laptop power consumption is an important attribute but perhaps not best described as a feature (although it may correspond to "battery life," which could be described as a feature).

[4] For example, clothing as a product category will have higher t, whereas most commodity products likely have smaller t, as their functional purpose is more important in determining product utility.

[5] Note that our model does not require that all consumers who purchase provide reviews. All that is needed for reviews to be informative is that a random set of purchasers provide reviews.

[6] We model a discrete rating system for tractability, although such a structure is consistent with many online services (e.g., Amazon) and any continuous system (or discrete system with greater granularity such as 1 to 5 "stars") can be mapped into a discrete system through choices of cutoffs. Unlike some prior work, we do not dichotomize ratings (H/L), because that loss of information prevents identification of certain key parameters—our method requires at least three levels.

[7] Note that this formulation assumes that there is no noise or bias in reviews. This should not affect the results, provided that review error is common across products.

[8] We thank the associate editor and an anonymous reviewer for suggesting these labels.

[9] Given that our focus in period 1 is to understand how consumers provide ratings, we assume that those who provide reviews provide ratings/reviews according to their realized utilities. Note that if (most) consumers do not truthfully provide ratings, then the whole value of a review system is limited. It is possible that sellers try to manipulate reviews in order to influence consumers' purchase decisions, and we will discuss the implications when reviews are fraudulent, or reviewers are not random in period 2 when we study the effects of reviews on later purchasers and subsequent ratings.

[10] As with all previous literature, we assume that the consumer's goal is to ensure that her purchase is satisfying (meaning the product provides utility that meets or exceeds reservation utility) rather than utility-maximizing, which requires the ability to sample and evaluate all products.

[11] When consumers are risk-averse, σ_v^2 will enter expected utility by lowering utility to $\mu - a \times \sigma_v^2$, where a reflects the level of risk aversion. Thus, $E[U(x,v)] = E[\tilde{v} - t \times \tilde{x}] = \mu - \frac{1}{2}t - A$, where $A = a \times (\sigma_v^2 + \sigma_x^2)$, the combined disutility due to risk aversion from quality and fit uncertainty. Note that $\sigma_x^2 = \frac{1}{12}$ (recall that x is uniformly distributed from 0 to 1). Consumers will make a purchase when $E[U(x,v)] \geq U_0$ (normalized to zero). One can see from these equations that the higher A is, the lower the expected utility, and the less likely consumers would make a purchase. Specifically, when consumers are risk-neutral, they will make a purchase when $\mu \geq \frac{1}{2}t$, but

when consumers are risk-averse (i.e., $A > 0$), the threshold to make a purchase becomes $\mu \geq \frac{1}{2}t + A$. Essentially, A matters only for expected utility and not realized utility and so only influences purchase decisions (and demand) but not consumers' ratings (which are conditional on purchasing and whether the realized utility exceeds reservation utility). Therefore, all of our analyses related to ratings hold regardless of A.

[12] Note that this definition of *experience* is similar in nature to Nelson's original classification using repair expenditures, in which a high ratio of repair expenditures over sales signals consumers' inability to assess product quality at the time of purchase.

[13] Consumers may still face uncertainty regarding whether the mean ratings indeed reflect true quality, as consumers may not be confident in trusting the mean ratings as the true value of the product, especially if random noise exists and the number of ratings is small. When consumers are risk-averse, then the expected utility will be lowered depending on how confident consumers are in trusting the mean ratings as a proxy of true quality. When consumers are not confident that $M_1(v)$ reflects true quality of the product, they will face some disutility due to risk aversion (A_2). When A_2 is not zero, it would affect a consumer's expected utility by lowering it by A_2, which again matters only for purchasing decisions. Everything else being equal, the lower A_2 is, the more likely consumers would make a purchase. Conditional on consumers making a purchase, A_2 would not matter for ratings, because ratings depend on consumption utility and reservation utility, neither of which involves A_2. Regardless, such uncertainty should go away as the number of ratings increases. Statistically, due to the law of large numbers, the higher the number of ratings, the closer the mean ratings are to the true quality of the product, and the lower the uncertainty.

[14] One may also interpret w as the fraction of consumers who are able to infer their fit measure based on existing consumer reviews. The more consumer reviews that are accumulated over time, the higher the fraction of consumers who may be able to learn their true fit measure.

[15] We focus on Amazon categories for the purposes of demonstrating the method acknowledging that categories may be heterogeneous. Such heterogeneity can be addressed by considering "subcategories" or conditioning the relevant statistics on measures that capture the heterogeneity. The appropriate approach will likely be dictated by the empirical application.

References

Archak N, Ghose A, Ipeirotis PG (2011) Deriving the pricing power of product features by mining consumer reviews. *Management Sci.* 57(8):1485–1509.

Brynjolfsson E, McAfee A (2011) *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy* (Digital Frontier Press, Lexington, MA).

Caves RE, Williamson PJ (1985) What is product differentiation, really? *J. Indust. Econom.* 34(2):113–132.

Chen Y, Xie J (2008) Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Sci.* 54(3):477–491.

Chen PY, Hong Y, Liu Y (2018) The value of multi-dimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Sci.* 64(10):4629–4647.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.

Clemons EK, Gao GG, Hitt LM (2006) When online reviews meet hyperdifferentiation: A study of the craft beer industry. *J. Management Inform. Systems* 23(2):149–171.

Darke PR, Chaiken S, Bohner G, Einwiller S, Erb HP, Hazlewood JD (1998) Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality Soc. Psych. Bull.* 24(11): 1205–1215.

Dellarocas C (2003) The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Sci.* 49(10):1407–1424.

Dellarocas C (2006) Strategic manipulation of Internet opinion forums: Implications for consumers and firms. *Management Sci.* 52(10):1577–1593.

Duan W, Gu B, Whinston AB (2008) The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *J. Retailing* 84(2):233–242.

Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inform. Systems Res.* 19(3):291–313.

Garvin DA (1984) What does "product quality" really mean? *Sloan Management Rev.* 26(1):25–43.

Ghose A, Ipeirotis P (2011) Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Trans. Knowledge Data Engrg.* 23(10):1498–1512.

Godes D, Silva JC (2012) Sequential and temporal dynamics of online opinion. *Marketing Sci.* 31(3):448–473.

Hotelling H (1929) Stability in competition. *Econom. J.* 39(153):41–57.

Hsieh Y-C, Chiu H-C, Chiang M-Y (2005) Maintaining a committed online customer: A study across search-experience-credence products. *J. Retailing* 81(1):75–82.

Hu N, Zhang J, Pavlou PA (2017) On selection biases of online product reviews. *MIS Quart.* 41(2):449–471.

Huang P, Lurie NH, Mitra S (2009) Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods. *J. Marketing* 73(2):55–69.

Huang S, Aral S, Hu YJ, Brynjolfsson E (2020) Social advertising eectiveness across products: a large-scale field experiment. *Marketing Sci.* 39(6):1142–1165.

Jiang Z, Benbasat I (2007) The effects of presentation formats and task complexity on online consumers' product understanding. *MIS Quart.* 31(3):475–500.

Jones Q, Ravid G, Rafaeli S (2004) Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Inform. Systems Res.* 15(2): 194–210.

Klein LR (1998) Evaluating the potential of interactive media through a new lens: Search vs. experience goods. *J. Bus. Res.* 41(3):195–203.

Kwark Y, Chen J, Raghunathan S (2014) Online product reviews: Implications for retailers and competing manufacturers. *Inform. Systems Res.* 25(1):93–110.

Laband DN (1991) An objective measure of search vs. experience goods. *Econom. Inquiry* 29(3):497–509.

Lancaster K (1966) A new theory of consumer demand. *J. Political Econom.* 74(2):132–157.

Li X, Wu L (2018) Herding and social media word-of-mouth: Evidence from Groupon. *MIS Quart.* 42(4):1331–1351.

Li X, Hitt LM (2010) Price effects in online product reviews: an analytical model and empirical analysis. *MIS Quart.* 34(4):809–831.

Li X, Hitt LM (2008) Self selection and information role of online product reviews. *Inform. Systems Res.* 19(4):456–474.

Li X, Hitt LM, Zhang ZJ (2011) Product reviews and competition in markets for repeat purchase products. *J. Management Inform. Systems* 27(4):9–42.

Liu Y (2006) Word of mouth for movies: Its dynamics and impact on box office revenue. *J. Marketing* 70(3):74–89.

Luca M, Zervas G (2016) Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Sci.* 62(12): 3412–3427.

Makadok R, Ross D (2013) Taking industry structuring seriously: A strategic perspective on product differentiation. *Strategic Management J.* 34(5):509–532.

McWilliams A, Siegel D (2001) Corporate social responsibility: A theory of the firm perspective. *Acad. Management Rev.* 26(1): 117–127.

Moe WW, Schweidel DA (2012) Online product opinions: Incidence, evaluation, and evolution. *Marketing Sci.* 31(3):372–386.

Muchnik L, Aral S, Taylor SJ (2013) Social influence bias: A randomized experiment. *Science* 341(6146):647–651.

Mudambi SM, Schuff D (2010) What makes a helpful online review? A study of customer reviews on Amazon. com. *MIS Quart.* 34(1): 185–200.

Nelson P (1970) Information and consumer behavior. *J. Political Econom.* 78(2):311–329.

Nelson P (1974) Advertising as information. *J. Political Econom.* 82(4): 729–754.

Nelson P (1981) Consumer information and advertising. Galatin M, Leiter RD, eds. Economics of Information (Springer, Dordrecht), 42–82.

Nisbett RE, Krantz DH, Jepson C, Kunda Z (1983) The use of statistical heuristics in everyday inductive reasoning. *Psych. Rev.* 90(4):339–363.

Park C, Lee TM (2009) Information direction, website reputation and eWOM effect: A moderating role of product type. *J. Bus. Res.* 62(1):61–67.

Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. *J. Political Econom.* 82(1):34–55.

Sanders WS, Nutakki GC, Nasraoui O (2016) Testing the application of warranting theory to online third party marketplaces: The effects of information uniqueness and product type. Gruzd A, Jacobson J, Mai P, Ruppert E, Murthy D, eds. *Proc. 7th 2016 Internat. Conf. Soc. Media Soc.* (ACM, New York), 1–7.

Senecal S, Nantel J (2004) The influence of online product recommendations on consumers' online choices. *J. Retailing* 80(2):159–169.

Shapiro C (1982) Consumer information, product quality, and seller reputation. *Bell J. Econom.* 13(1):20–35.

Simon HA (1982) *Models of Bounded Rationality: Empirically Grounded Economic Reason* (MIT Press, Cambridge, MA).

Stigler G (1961) The economics of information. *J. Political Econom.* 69(3):213–225.

Stock JH, Watson MW (2003) *Introduction to Econometrics* (Addison Wesley, Boston).

Sun M (2012) How does variance of product ratings matter? *Management Sci.* 58(4):696–707.

Theil H (1971) *Principles of Econometrics* (Wiley, New York).

Vandenbosch M, Weinberg C (1995) Product and price competition in a two-dimensional vertical differentiation model. *Marketing Sci.* 14(2):224–249.

Villas-Boas JM (2004) Consumer learning, brand loyalty, and competition. *Marketing Sci.* 23(1):134–145.

Villas-Boas JM (2006) Dynamic competition with experience goods. *J. Econom. Management Strategy* 15(1):37–66.

Wang C, Zhang X, Hann IH (2018) Socially nudged: A quasi-experimental study of friends' social influence in online product ratings. *Inform. Systems. Res.* 29(3):641–655.

Weathers D, Sharma S, Wood SL (2007) Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods. *J. Retailing* 83(4):393–401.

Zhai Q, Cao X, Mokhtarian PL, Zhen F (2017) The interactions between e-shopping and store shopping in the shopping process for search goods and experience goods. *Transportation* 44(5): 885–904.

Zhu F, Zhang X (2010) Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *J. Marketing* 74(2):133–148.